Exact Bayesian Learning of Ancestor Relations in Bayesian Networks
Supplementary Material

Yetian Chen yetianc@iastate.edu
Lingjian Meng lmeng@iastate.edu
Jin Tian jtian@iastate.edu
Department of Computer Science
Iowa State University
Ames, IA 50011, USA

Abstract

This technical report provides supplementary material for the paper "Exact Bayesian Learning of Ancestor Relations in Bayesian Networks". Here we extend our algorithm to compute the exact posterior of any $s \rightsquigarrow p \rightsquigarrow t$ relation, i.e., a directed path from s to t via p, in $O(n7^{n-2})$ time and $O(4^{n-2})$ space.

1. Computing Posteriors of $s \rightsquigarrow p \rightsquigarrow t$ Relations

1.1 Algorithm

The problem is to evaluate whether there is a directed path from s to t via p. Similarly, we would like to compute the joint probability $P(s \rightsquigarrow p \rightsquigarrow t, D)$ by

$$P(s \rightsquigarrow p \rightsquigarrow t, D) = \sum_{G: s \rightsquigarrow p \rightsquigarrow t \in G} \prod_{i \in V} B_i(Pa_i^G). \tag{1}$$

For any T, R, S such that $p \in T \subset R \subseteq S \subseteq V$, $s \in R - T$, let $\mathcal{G}_{s,p}(S,R,T)$ denote the set of all possible DAGs over S such that R are the set of all descendants of s (including s) and T are the set of all descendants of p (including p) in G_S. That is, $G_S \in \mathcal{G}_{s,p}(S,R,T)$ if and only if $de_{G_S}(s) = R$ and $de_{G_S}(p) = T$. We then define

$$H_{s,p}(S,R,T) \equiv \sum_{G_S \in \mathcal{G}_{s,p}(S,R,T)} \prod_{i \in S} B_i(Pa_i^{G_S}). \tag{2}$$

Then we have

Lemma 1

$$P(s \rightsquigarrow p \rightsquigarrow t, D) = \sum_{T,R: \{p,t\} \subseteq T \subset R \subseteq V, s \in R - T} H_{s,p}(V,R,T). \tag{3}$$

Proof . Let $\mathcal{G}_{s \rightsquigarrow p \rightsquigarrow t} = \{G : s \rightsquigarrow p \rightsquigarrow t \in G\}$, namely the set of all possible DAGs over V that contains a $s \rightsquigarrow p \rightsquigarrow t$. Then we have $\mathcal{G}_{s \rightsquigarrow p \rightsquigarrow t} = \cup_{T,R: \{p,t\} \subseteq T \subset R \subseteq V, s \in R-T} \mathcal{G}_{s,p}(V,R,T)$. Further, for any $T_1 \neq T_2$ or $R_1 \neq R_2$, we have $\mathcal{G}_{s,p}(V,R_1,T_1) \cap \mathcal{G}_{s,p}(V,R_2,T_2) = \emptyset$. This means $\mathcal{G}_{s,p}(V,R,T)$ for all T, R such that $p \in T \subset R \subseteq V$, $s \in R - T$ form a partition of the set $\mathcal{G}_{s \rightsquigarrow p \rightsquigarrow t}$. Thus,

$$
\begin{aligned}
P(s \rightsquigarrow p \rightsquigarrow t, D) &= \sum_{G \in \mathcal{G}_{s \rightsquigarrow p \rightsquigarrow t}} \prod_{i \in V} B_i(Pa_i^G) = \sum_{\substack{T,R:\{p,t\} \subseteq T \subset R \subseteq V \\ s \in R-T}} \sum_{G \in \mathcal{G}_{s,p}(V,R,T)} \prod_{i \in V} B_i(Pa_i^G) \\
&= \sum_{T,R:\{p,t\} \subseteq T \subset R \subseteq V, s \in R-T} H_{s,p}(V,R,T).
\end{aligned} \tag{4}
$$

■

If we have all $H_{s,p}(S,R,T)$ computed, we can compute Eq. (4) in $\sum_{|R|=1}^{n} \left[\binom{n-3}{|R|-3} \sum_{|T|=2}^{|R|-1} \binom{|R|-3}{|T|-2} \right] = O(3^{n-3})$ time.

Now we can show that $H_{s,p}(S,R,T)$ for all T, R, S such that $p \in T \subset R \subseteq S \subseteq V$ and $s \in R-T$ can be computed recursively. These $H_{s,p}(S,R,T)$'s can be divided into two cases: $T = \{p\}$ and $T \neq \{p\}$.

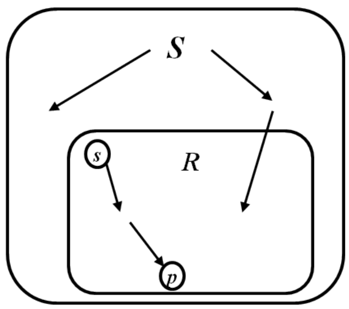


Figure 1: Case 1: $T = \{p\}$.

Case 1: $T = \{p\}$.

In this case, p is a sink in G_S (see Figure 1) and its parent set must include a least one node in $R - \{p\}$ to make it a descendant of s. For nodes in $S - \{p\}$, we have summation over $\mathcal{G}_s(S - \{p\}, R - \{p\})$, i.e., the set of DAGs over $S - \{p\}$ s.t. $R - \{p\}$ are the set of descendants of s in $G_{S-\{p\}}$. Then we have

$$
\begin{aligned}
H_{s,p}(S,R,\{p\}) &= [\sum_{\substack{Pa_p \subseteq S-\{p\} \\ Pa_p \cap R-\{p\} \neq \emptyset}} B_p(Pa_p)][\sum_{G_{S-\{p\}} \in \mathcal{G}_s(S-\{p\},R-\{p\})} \prod_{i \in S-\{p\}} B_i(Pa_i^{G_{S-\{p\}}})] \\
&= [\sum_{Pa_p \subseteq S-\{p\}} B_p(Pa_p) - \sum_{Pa_p \subseteq S-R} B_p(Pa_p)] H_s(S-\{p\}, R-\{p\}) \\
&= [A_p(S-\{p\}) - A_p(S-R)] H_s(S-\{p\}, R-\{p\}) \\
&= [AA(S-\{p\},\{p\}) - AA(S-R,\{p\})] H_s(S-\{p\}, R-\{p\}).
\end{aligned} \tag{5}
$$

Case 2: $T \neq \{p\}$.

For any $W \subseteq S - \{s,p\}$, let $\mathcal{G}_{s,p}(S,R,T,W)$ denote the set of DAGs in $\mathcal{G}_{s,p}(S,R,T)$ such that all nodes in W are (must be) sinks. [1] Then we define

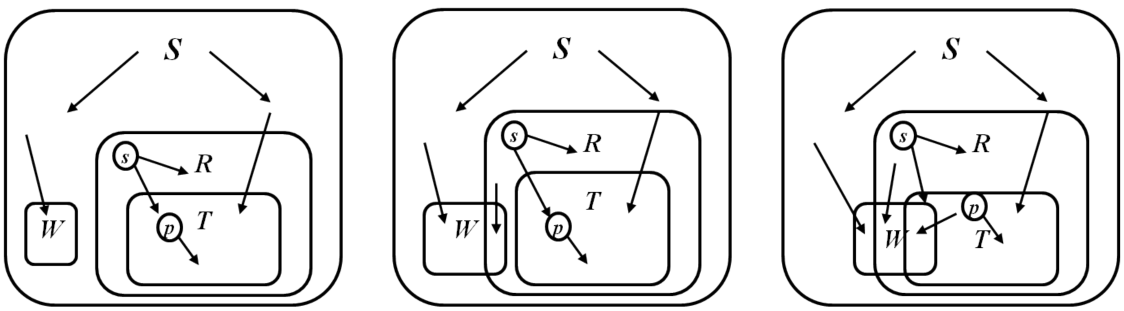
1. Again, W may not include all the sinks in G_S. Some nodes in $S - W$ could be sinks.

$$F_{s,p}(S,R,T,W) \equiv \sum_{G_S \in \mathcal{G}_{s,p}(S,R,T,W)} \prod_{i \in S} B_i(Pa_i^{G_S}). \tag{6}$$

Similarly, by weighted inclusion-exclusion principle,

$$\begin{aligned} H_{s,p}(S,R,T) &= \sum_{k=1}^{|S|-2} (-1)^{k+1} \sum_{W \subseteq S-\{s,p\}, |W|=k} \sum_{G_S \in \mathcal{G}_{s,p}(S,R,T,W)} \prod_{i \in S} B_i(Pa_i) \\ &= \sum_{k=1}^{|S|-2} (-1)^{k+1} \sum_{W \subseteq S-\{s,p\}, |W|=k} F_{s,p}(S,R,T,W). \end{aligned} \tag{7}$$

$F_{s,p}(S,R,T,W)$ and $H_{s,p}(S,R,T)$ can be computed recursively. There are three sub-cases (see Figure 2).



(a) $W \cap R = \emptyset$ (b) $W \cap R \neq \emptyset$ and $W \cap T = \emptyset$ (c) $W \cap T \neq \emptyset$

Figure 2: Three different cases when computing $F_{s,p}(S,R,T,W)$.

Sub-case 1: $W \cap R = \emptyset$.

We can compute the summation for W and $S-W$ separately (see Figure 2(a)). We have

$$\begin{aligned} &F_{s,p}(S,R,T,W) \\ &= \Big[\prod_{j \in W} \sum_{Pa_j \subseteq (S-R-W)} B_j(Pa_j)\Big]\Big[\sum_{G_{S-W} \in \mathcal{G}_{s,p}(S-W,R,T)} \prod_{i \in S-W} B_i(Pa_i^{G_{S-W}})\Big] \\ &= \prod_{j \in W} A_j(S-R-W) H_{s,p}(S-W,R,T) = AA(S-R-W,W) H_{s,p}(S-W,R-W,T-W) \end{aligned} \tag{8}$$

(because $R-W=R$ and $T-W=T$ in this case).

Sub-case 2: $W \cap R \neq \emptyset$ **and** $W \cap T = \emptyset$.

In this case, nodes in $W-R$, $W \cap R$, and $S-W$ should be handled separately (see Figure 2(b)). Nodes in $W-R$ can only select parents from $S-R-W$. Any node in $W \cap R$ can select parents from $S-W-T$. In addition, at least one node from $R-T-W$ must be included in its parent set to guarantee that it is a descendant of s. For nodes in $S-W$, we have summation over $\mathcal{G}_{s,p}(S-W,R-W,T)$. Then we have

$$F_{s,p}(S,R,T,W) = [\prod_{j\in W-R}\sum_{Pa_j\subseteq(S-R-W)} B_j(Pa_j)][\prod_{j\in W\cap R}\sum_{\substack{Pa_j\subseteq(S-W-T)\\ Pa_j\cap(R-T-W)\neq\emptyset}} B_j(Pa_j)] \tag{9}$$

$$[\sum_{\substack{G_{S-W}\in\\ \mathcal{G}_{s,p}(S-W,R-W,T)}}\prod_{i\in S-W} B_i(Pa_i^{G_{S-W}})]$$

$$= \prod_{j\in W-R} A_j(S-R-W)\left\{\prod_{j\in W\cap R}[A_j(S-W-T)-A_j(S-W-R)]\right\} H_{s,p}(S-W,R-W,T)$$

$$= AA(S-W-R,W-R)\left\{\prod_{j\in W\cap R}[A_j(S-W-T)-A_j(S-W-R)]\right\}$$

$H_{s,p}(S-W,R-W,T-W)$ (because $T-W=\emptyset$ in this case).

Sub-case 3: $W\cap T\neq\emptyset$.

In this case, nodes in $W-R$, $W\cap(R-T)$, $W\cap T$, and $S-W$ should be handled separately (see Figure 2(c)). Nodes in $W-R$ can only select parents from $S-R-W$. Any node in $W\cap(R-T)$ can select parents from $S-W-T$. In addition, at least one node from $R-T-W$ must be included in its parent set to guarantee that it is a descendant of s. Nodes in $W\cap T$ can select parents from $S-W$ and at least one node as its parent from $T-W$ to make it a descendant of p. For nodes in $S-W$, we have summation over $\mathcal{G}_{s,p}(S-W,R-W,T-W)$. Then we have

$$F_{s,p}(S,R,T,W)$$

$$= [\prod_{j\in W-R}\sum_{Pa_j\subseteq(S-R-W)} B_j(Pa_j)][\prod_{j\in W\cap(R-T)}\sum_{\substack{Pa_j\subseteq(S-W-T)\\ Pa_j\cap(R-T-W)\neq\emptyset}} B_j(Pa_j)]$$

$$[\prod_{j\in W\cap T}\sum_{\substack{Pa_j\subseteq(S-W)\\ Pa_j\cap(T-W)\neq\emptyset}} B_j(Pa_j)][\sum_{\substack{G_{S-W}\in\\ \mathcal{G}_{s,p}(S-W,R-W,T-W)}}\prod_{i\in S-W} B_i(Pa_i^{G_{S-W}})]$$

$$= \prod_{j\in W-R} A_j(S-R-W)\left\{\prod_{j\in W\cap(R-T)}[A_j(S-W-T)-A_j(S-W-R)]\right\} \tag{10}$$

$$\left\{\prod_{j\in W\cap T}[A_j(S-W)-A_j(S-W-T)]\right\} H_{s,p}(S-W,R-W,T-W)$$

$$= AA(S-W-R,W-R)\left\{\prod_{j\in W\cap(R-T)}[A_j(S-W-T)-A_j(S-W-R)]\right\}$$

$$\left\{\prod_{j\in W\cap T}[A_j(S-W)-A_j(S-W-T)]\right\} H_{s,p}(S-W,R-W,T-W).$$

For ease of exposition, for all S, R, T, W such that $\{p\} \subset T \subset R \subseteq S \subseteq V$, $s \in R - T$ and $W \subseteq S - \{s, p\}$, define function $\mathcal{A}_{s,p}(S, R, T, W)$ as follows:

$$
\begin{aligned}
&\text{If } W \cap R = \emptyset, && (11)\\
&\mathcal{A}_{s,p}(S,R,T,W) \equiv AA(S-R-W, W);\\
&\text{If } W \cap R \neq \emptyset \text{ and } W \cap T = \emptyset,\\
&\mathcal{A}_{s,p}(S,R,T,W) \equiv AA(S-W-R, W-R)\left\{\prod_{j \in W \cap R}[A_j(S-W-T) - A_j(S-W-R)]\right\};\\
&\text{If } W \cap T \neq \emptyset,\\
&\mathcal{A}_{s,p}(S,R,T,W) \equiv AA(S-W-R, W-R)\left\{\prod_{j \in W \cap (R-T)}[A_j(S-W-T) - A_j(S-W-R)]\right\}\\
&\left\{\prod_{j \in W \cap T}[A_j(S-W) - A_j(S-W-T)]\right\}.
\end{aligned}
$$

Now $F_{s,p}(S, R, T, W)$ can be neatly written as

$$
F_{s,p}(S,R,T,W) = \mathcal{A}_{s,p}(S,R,T,W) H_{s,p}(S-W, R-W, T-W). \tag{12}
$$

Then we have a recursive formula for computing $H_{s,p}(S, R, T)$,

$$
H_{s,p}(S,R,T) = \sum_{k=1}^{|S|-2} (-1)^{k+1} \sum_{W \subseteq S-\{s,p\}, |W|=k} \mathcal{A}_{s,p}(S,R,T,W) H_{s,p}(S-W, R-W, T-W). \tag{13}
$$

And finally, we arrive the following recursive scheme for computing $H_{s,p}(S, R, T)$ for all T, R, S such that $p \in T \subset R \subseteq S \subseteq V$ and $s \in R - T$.

Theorem 2

$$
\begin{aligned}
&H_{s,p}(S,R,\{p\}) = [AA(S-\{p\},\{p\}) - AA(S-R,\{p\})] H_s(S-\{p\}, R-\{p\})\\
&\textit{for all } \{s,p\} \subseteq R \subseteq S \subseteq V,\\
&H_{s,p}(S,R,T) = \sum_{k=1}^{|S|-2} (-1)^{k+1} \sum_{W \subseteq S-\{s,p\}, |W|=k} \mathcal{A}_{s,p}(S,R,T,W) H_{s,p}(S-W, R-W, T-W)\\
&\textit{for all } T, R, S \textit{ s.t } \{p\} \subset T \subset R \subseteq S \subseteq V \textit{ and } s \in R - T.
\end{aligned} \tag{14}
$$

Note that all $H_s(S - \{p\}, R - \{p\})$'s can be computed recursively using **Theorem 2**.

1.2 Time and Space Complexity

Computing $H_{s,p}(S,R,\{p\})$ and $H_{s,p}(S,R,T)$ dominates the total computation time. Given all $H_s(S-\{p\},R-\{p\})$'s pre-computed, $H_{s,p}(S,R,\{p\})$ for all $\{s,p\}\subseteq R\subseteq S\subseteq V$ can be computed in $\sum_{|S|=2}^{n}\binom{n-2}{|S|-2}\sum_{|R|=2}^{|S|}\binom{|S|-2}{|R|-2}=O(3^{n-2})$ time. All other $H_{s,p}(S,R,T)$'s can be computed in

$$\begin{aligned}
&\sum_{|S|=3}^{n}\binom{n-2}{|S|-2}\left\{\sum_{|R|=3}^{|S|}\binom{|S|-2}{|R|-2}\left[\sum_{|T|=2}^{|R|-1}\binom{|R|-2}{|T|-1}|S|\cdot 2^{|S|-2}\right]\right\} \qquad (15)\\
&=\sum_{|S|=3}^{n}\binom{n-2}{|S|-2}\left\{\sum_{|R|=3}^{|S|}\binom{|S|-2}{|R|-2}|S|\cdot 2^{|S|+|R|-4}\right\}=\sum_{|S|=3}^{n}\binom{n-2}{|S|-2}\left[|S|\cdot 2^{|S|-2}\cdot 3^{|S|-2}\right]\\
&=\sum_{|S|=3}^{n}\binom{n-2}{|S|-2}\left[|S|\cdot 6^{|S|-2}\right]<n7^{n-2}.
\end{aligned}$$

Thus, the total computation time is $O(n7^{n-2})$. The space complexity is dominated by $H_{s,p}(S,R,T)$, which is

$$\begin{aligned}
&\sum_{|S|=2}^{n}\binom{n-2}{|S|-2}\left\{\sum_{|R|=2}^{|S|}\binom{|S|-2}{|R|-2}\left[\sum_{|T|=1}^{|R|-1}\binom{|R|-2}{|T|-1}\right]\right\}\\
&=\sum_{|S|=2}^{n}\binom{n-2}{|S|-2}\left\{\sum_{|R|=2}^{|S|}\binom{|S|-2}{|R|-2}2^{|R|-2}\right\}=\sum_{|S|=3}^{n}\binom{n-2}{|S|-2}3^{|S|-2}=4^{n-2}.
\end{aligned} \qquad (16)$$

Thus, the total space requirement is $O(4^{n-2}+3^n)$. Thus, we have the following theorem.

Theorem 3 *The posterior probability of any $s \rightsquigarrow p \rightsquigarrow t$ relation can be computed in $O(n7^{n-2})$ time and $O(4^{n-2}+3^n)$ space.*