Bounding the False Discovery Rate in Local Bayesian Network Learning

Supplemental Material

Theorem 2. *Let* $\mathbf{EN}_T, \mathbf{EN}_i$ *be any subsets of variables such that* $\mathbf{EN}_T \supseteq \mathbf{N}_T$ *and* $\mathbf{EN}_i \supseteq \mathbf{N}_i$*. Then,*

$$V_i \notin \mathbf{N}_T \Leftrightarrow \exists \mathbf{Z}_k \in 2^{\mathbf{EN}_T \setminus \{V_i, T\}} \cup 2^{\mathbf{EN}_i \setminus \{V_i, T\}}, \quad \text{s.t. } I(V_i; T | \mathbf{Z_k})$$

Proof. If a certificate of exclusion exists in $\mathbf{EN}_T$ or in $\mathbf{EN}_i$, then by Theorem 1, $V_i \notin \mathbf{N}_T$. Conversely, let $V_i \notin \mathbf{N}_T$ and let $\mathcal{G}$ be the graph of any network faithful to the data distribution. Let $Pa_T \subseteq \mathbf{N}_T \setminus \{V_i, T\} \subseteq \mathbf{EN}_T \setminus \{V_i, T\}$ be the parents of T in $\mathcal{G}$, and $Pa_i \subseteq \mathbf{N}_i \setminus \{V_i, T\} \subseteq \mathbf{EN}_i \setminus \{V_i, T\}$ be the parents of V_i in $\mathcal{G}$ (note $Pa_T \subseteq \mathbf{N}_T \setminus \{V_i, T\}$, because we assumed $V_i \notin \mathbf{N}_T$; similarly for Pa_i). If V_i is a non-descendant of T, then $I(V_i; T | Pa_T)$ by the Markov Condition. Similarly, if T is a non-descendant of V_i, then $I(V_i; T | Pa_i)$. Either way, there exist a subset in $\mathbf{EN}_T$ or in $\mathbf{EN}_i$ such that the independence holds. One of the two cases has to hold because the network is acyclic. □