

Algorithm HITON-PC (without “symmetry correction”)

Input: dataset with predictive variables (e.g., genes) \mathbf{V} , including a phenotypic response variable T

Output: a Markov boundary \mathbf{M} of T

1. Initialize \mathbf{M} with an empty set
2. Initialize the set of eligible variables $\mathbf{E} \leftarrow \mathbf{V} \setminus \{T\}$
3. Sort in descending order the variables in \mathbf{E} according to their pairwise association with response variable T
4. Remove from \mathbf{E} all variables X with zero association with T , i.e. when $T \perp X$
5. Repeat
6. $X \leftarrow$ first variable in \mathbf{E}
7. Add X to \mathbf{M} and remove it from \mathbf{E}
8. If $\exists \mathbf{Z} \subseteq \mathbf{M} \setminus \{X\}$, such that $T \perp X \mid \mathbf{Z}$, remove X from \mathbf{M}
9. Until \mathbf{E} is empty
10. For each $X \in \mathbf{M}$
11. If $\exists \mathbf{Z} \subseteq \mathbf{M} \setminus \{X\}$, such that $T \perp X \mid \mathbf{Z}$, remove X from \mathbf{M}
12. Output \mathbf{M}

Figure S2: HITON-PC algorithm. More details about the algorithm and characterization of distributions where it can correctly identify a Markov boundary are given in [1,2].

References

1. Aliferis CF, Statnikov A, Tsamardinos I, Mani S, Koutsoukos XD (2010) Local Causal and Markov Blanket Induction for Causal Discovery and Feature Selection for Classification. Part II: Analysis and Extensions. Journal of Machine Learning Research 11: 235-284.
2. Aliferis CF, Statnikov A, Tsamardinos I, Mani S, Koutsoukos XD (2010) Local Causal and Markov Blanket Induction for Causal Discovery and Feature Selection for Classification. Part I: Algorithms and Empirical Evaluation. Journal of Machine Learning Research 11: 171-234.