Noname manuscript No.
(will be inserted by the editor)

Weak recursively simplicial graphs

First Author · Second Author · Third Author

Received: date / Accepted: date

1 Preliminary

Definition 1 Let $F = (V, E)$ and $F' = (V', E')$ be two graphs. If $V' \subseteq V$ and $E' \subseteq E$, then F' is a **subgraph** of F (and F is a **supergraph** of F'), written as $F' \subseteq F$. If $F' \subseteq F$ and F' contains all the edges $xy \in E$ with $x, y \in V'$, then F' is an **induced subgraph** of F, written as $F' = F[V']$.

Definition 2 A **clique** is a subset of nodes in an undirected graph where every two distinct nodes are adjacent.

Definition 3 A **simplicial node** in an undirected graph is a node whose neighbours form a clique.

Definition 4 A graph $F = (V, E)$ is **recursively simplicial** if it contains a simplicial node v_i and the induced subgraph $F[V \setminus \{v_i\}]$ is recursively simplicial.

Definition 5 A **simplicial node ordering** of a recursively simplicial graph $F = (V, E)$ is a sequence of nodes $\{v_1, \ldots, v_n\}$, where v_i is a simplicial node in the induced subgraph $F[V \setminus \{v_1, \ldots, v_{i-1}\}]$.

Definition 6 An **m-cycle** in an undirected graph is a sequence of nodes $\{v_1, \ldots, v_{m+1}\}$ where $v_1 = v_{m+1}$ and all the other nodes are distinct.

Definition 7 A graph is **chordal** if each m-cycle for $m \geq 4$ has a chord.

F. Author
first address
Tel.: +123-45-678910
Fax: +123-45-678910
E-mail: fauthor@example.com

S. Author
second address

Definition 8 Let $< G, P >$ be a graphical model over a set $V = \{v_1, \ldots, v_n\}$ of n random variables. The **Markvov blanket** of a variable v_i in G, denoted by B_i^G, is a subset of variables such that $v_i \perp\!\!\!\perp_P V \setminus \{B_i^G \cup \{v_i\}\} \mid B_i^G$.

In a DAG, B_i consists of v_i's parents, children and spouses (i.e., children's other parents). In a undirected graph (UG), B_i consists of the (distance 1) neighbours of v_i. We use B_V to denote the family of Markov blankets $\{B_1, \ldots, B_n\}$ over V.

Definition 9 A **hybrid graph** is a graph consisting of directed and undirected edges.

Definition 10 The **skeleton** of a hybrid graph is the undirected graph obtained by dropping directions of all directed edges.

Definition 11 The **moral graph** of a DAG is the skeleton of the hybrid graph obtained by adding undirected edges between all pairs of non-adjacent parents that have a common child in the DAG.

The following are some well-known properties of chordal graphs (citation!!!).

Proposition 1 *The following properties of an undirected graph F are equivalent:*

1. *F is chordal.*
2. *F is recursively simplicial.*
3. *F can be oriented into a DAG G s.t. the moral graph of G is F.*
4. *F can be oriented into a DAG G s.t. F and G imply the same conditional independencies.*

As a consequence of 1 and 4, F can be oriented into a DAG G s.t. $B_X^F = B_X^G$. The following proposition states that there exists non-chordal graphs, who can be oriented into DAGs with the same Markov blanket families.

Proposition 2 *Let $\mathcal{F}$ be the set of chordal graphs $F = (V, E)$ and $\mathcal{B}$ be the set of Markov blanket families of DAGs over V. Then there exists an injective but non-surjective function $\phi : \mathcal{F} \to \mathcal{B}$.*

Proof Let F_1, F_2 be two identical chordal graphs. It implies $B_V^{F_1} = B_V^{F_2}$. Since F_i is a chordal, it can be oriented into a DAG G_i s.t. $B_V^{F_i} = B_V^{G_i}, \forall i \in \{1, 2\}$. It follows that $B_V^{G_1} = B_V^{G_2}$, so ϕ is an injective function.

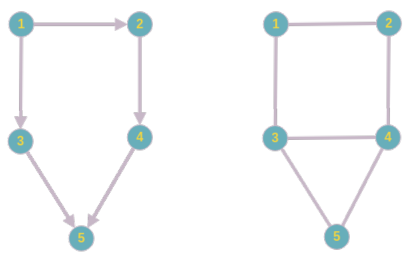

Fig. 1: An example of a DAG whose moral graph is non-chordal.

For any given DAG G, it has a unique moral graph F s.t. $B_V^G = B_V^F$. Figure 1 is an example of a DAG and its moral graph, which is non-choral. Hence, ϕ is non-surjective. □

2 Weak recursively simplicial graphs

Definition 12 A **collider** in a hybrid graph is a node with at least two parents.

Definition 13 A DAG G is a **consistent extension** of a hybrid graph H if G and H have the same skeleton and the same set of colliders.

Let $F = (V, E)$ be a graph over n nodes. Let $S = \{S_1, \ldots, S_m\}$ be a set of disjoint subsets of V s.t. $\cup_{i=1}^m S_i = V$. Let $U = \{U_1, \ldots, U_m\}$ be a set of disjoint subsets of E, where $U_i \subseteq \{v_j v_k \mid \forall v_j, v_k \in N_l, \forall v_l \in S_i\}$. Two extremes of U_i are the empty set and the edges between all pairs of neighbours of v_l for all $v_l \in S_i$. We use $O = \{(S_1, U_1), \ldots, (S_m, U_m)\}$ to denote the pairs. Next we introduce a subset of undirected graphs which are not as strong as recusrively simplicial.

Definition 14 A graph $F = (V, E)$ is **weak recursively simplicial (WRS)** if there exits $O = \{(S_1, U_1), \ldots, (S_m, U_m)\}$ such that S_1 is the set of simplicial nodes in F and S_i is the set of simplicial nodes in the subgraph obtained by removing from F the pairs $\{O_1, \ldots, O_{i-1}\}$ for $i \in [2, m]$.

The reason such a graph is named weak recursively simplicial is because if we only remove the simplicial nodes from F, the induced subgraph over the remaining nodes may or may not have any simplicial nodes. So none or some edges between neighbours of simplicial nodes may need to be removed as well to retain the recursion. By definition, a chordal graph is also weak recursively simplicial by letting $U_i = \emptyset$. The converse, however, is not true. The graph on the right of Figure 1 is a WRS graph. It has $O_1 = \{S_1 = \{5\}, U_1 = \{3-4\}\}$. The subgraph obtained by removing O_1 is a chordal graph that is also WRS. Therefore, the original graph is WRS.

Proposition 3 *Let $G = (V, E)$ be a DAG and F be the moral graph of G. Then F is weak recursively simplicial.*

Proof (Base case) $G(1) = F(1)$ is a single node graph, so $F(1)$ is WRS. (Inductive hypothesis) Assuming the moral graph $F(n)$ of any DAG $G(n)$ with n nodes is WRS. (Inductive step) Adding the node v_{n+1} in $G(n)$ as a new leaf produces another DAG $G(n+1)$ with $n+1$ nodes. Since v_{n+1} is a leaf in $G(n+1)$, it becomes a simplicial node in $F(n+1)$ after moralization. In addition, extra edges may be introduced to connect parents of v_{n+1}. Since v_{n+1} is a simplicial node, it can be removed from $F(n+1)$ together with the extra edges introduced. Hence, we get back to the same moral graph $F(n)$ that is assumed to be WRS. Therefore, $F(n+1)$ is also WRS. □

Proposition 3 suggests that the moral graph of any DAG is WRS. To prove there is a one-to-one correspondance between moral graphs and WRS graphs, we want to show that the converse is also true. That is, a WRS graph F can be oriented into a DAG G (possibly with some edge removal) such that the moral graph of G is identical to F.

Proposition 4 *Let $F = (V, E)$ be a weak recursively simplicial graph. Then F is the moral graph of a DAG.*

Proof (Base case) $F(1)$ is a single node WRS graph, and also the moral graph of the single node DAG $G(1)$. (Inductive hypothesis) Assuming any WRS graph $F(n)$ with n nodes is the moral graph of a DAG $G(n)$. (Inductive step) Adding v_{n+1} into $F(n)$ as a simplical node, so the resulting graph $F(n+1)$ is WRS. Since v_{n+1} is a simplicial node, there exists a DAG $G(n+1)$, in which v_{n+1} is a leaf. Therefore, the moral graph of $G(n+1)$ is $F(n+1)$. □

Corollary 1 *An undirected graph is weak recursively simplicial if and only if it is the moral graph of a DAG.*

Proof The corollary follows from the Proposition 3 and Proposition 4. □

Corollary 2 *Let $\mathcal{F} = \{F = (V, E) \mid F$ is weak recursively simplicial$\}$ be the set of weak recursively simplicial graphs over V and $\mathcal{B} = \{B_V^G \mid$ for all DAG G over $V\}$ be the set of Markov blanket families of any DAG over V. Then $|\mathcal{F}| = |\mathcal{B}|$.*

Proof It is straightforward that there is a one-to-one correspondance between $\mathcal{B}$ and moral graphs. Hence, Corollary 1 implies that $\mathcal{F}$ has a one-to-one correspondance with $\mathcal{B}$, so equal cardinality. □

Next, we present a backtracking algorithm for checking whether or not a given graph $F = (V, E)$ is WRS. If it is, the algorithm will find a DAG G, the moral graph of whom is F. Notice that this algorithm may produce a hybrid graph, which always has a consistent DAG extension according to Proposition 4. Hence, we use the algorithm presented in [Dor and Tarsi, 1992] to get it.

Algorithm 1 Backtracking algorithm to test WRS of graphs

Require: F is an undirected graph obtained by connecting all $v_j \in B_i$ with $v_j, \forall v_j \in V$

```
function WRS(F)
    if F is chordal then return TRUE
    end if
    Find all the set S of all simplicial nodes in F
    if S ≠ ∅ then
        Find E = {v_j v_k | ∀v_j, v_k ∈ N_i, ∀v_i ∈ S}
        backup = F
        for each U ⊆ E do                          ▷ for each subset of E
            F = backup
            F = F − S − U                          ▷ Remove simplicial nodes S and edges U
            if WRS(F) = TRUE then return TRUE
            end if
        end for
    end if
    return FALSE
end function
```

[Verma and Pearl, 1992] did similar work but for finding consistent DAG with the set of conditional independencies. It maybe equivalent to MBs?

Table 1: Comparison between the number of chordal graphs , the number of weak recursively simplicial graphs, the number of undirected graphs and the number of Markov equivalent classes.

# nodes	# chordal graphs	# weak r.s. graphs	# undirected graphs	# Markov equivalent classes
1	1	1	1	1
2	2	2	2	2
3	8	8	8	11
4	61	61	64	185
5	822	882	1024	8782
6	18154		32768	1067825
7	617675		2097152	312510571
8	30888596		268435456	212133402500
9	2192816760		68719476736	326266056291213
10	215488096587 (2×10^{11})		35184372088832 (3×10^{13})	118902054495975141 (1×10^{17})

3 Some random notes

Lloyd's conjecture: knowing a graph is wrs, adding an edge to obtain a supergraph, perhaps it is efficient to check if the supergraph is wrs.

Proof because remove the added edge, then we obtain the original graph which we know is wrs. however, we don't remove a random edge when checking wrs unless the edge connects to a simplicial node. what if the edge is not connected with a sim node? if we know the original graph is wrs and we know the set of sim nodes in the recursion and the set of edges to delete, then we could easily check if the new edge is connected with any one of the sim node, if it is then good. if not, then we could check if the new edge is connected with two neighbours of a sim node, if it is then good. if a graph is wrs, then it eventually will diminish, so the new edge must appear somewhere in the recursion to either stop the recursion or don't stop it.

Corollary 3 *DAGs in the same Markov equivalent class produce the same Markov blanket sets B_X.*

Proof If two DAGs G_1 and G_2 are Markov equivalent, they have the same skeleton and the same set of colliders. This implies $B_i^{G_1} = B_i^{G_2}, \forall X_i \in X$. □

Notice that two Markov equivalent classes could entail the same B_X. For example...

Corollary 4 $|\{\textit{chordal graphs}\}| \leq |B_X| \leq |\{\textit{Markov equivalent classes}\}|$.

Counting labelled chordal graphs [Wormald, 1985], counting Markov equivalent classes (assymptotic ratio of around 0.27 to DAGs) [Gillispie and Perlman, 2001].

Proposition 5 *Let G be a DAG and F be the moral graph of G. If a node x is a leaf in G, then it must be a simplicial node in F.*

Proof If x is a leaf in G, it has only parents, which form a clique after moralization. By definition, x is a simplicial node in F. □

Corollary 5 *Let G be a DAG and F be the moral graph of G. Then F must have at least one simplicial node.*

Proof Since each DAG has at least one leaf, by Proposition 5 F have at least one simplicial node. ☐

Corollary 6 *Let G be a DAG and F be the moral graph of G. If a node x is not a simplicial node in F, then it must not be a leaf in G.*

Proposition 6 *Let G be a DAG and F be the moral graph of G. Let S^1 be the set of simplicial nodes in F and F_1 be the induced subgraph of F over $X \setminus S^1$. Then there must exist at least one simplicial node after removing from F' all the edges between $N(X_i), \forall X_i \in S^1$.*

Proof Let F_1' be the result of removing from F_1 all the edges between $N(X_i), \forall X_i \in S^1$. The corresponding directed graph G' of F_1' must be a subgraph of the DAG G, so also acyclic. Assuming F_1' has no simplicial nodes, by Corollary 6 G' has no leaf, which is a contradiction. ☐

Here are some issues worth discussing:

1. Application: the backtracking algorithm can now be applied when learning MBs in paralle. What if there are conflicts between two MBs, which one should give up? Need to estimate uncertainty?
2. Simplicial nodes in the first step always contain the leaves.
3. Those nodes that become simplicial in the next step without having to delete any edges contain the leaves in the next step.
4. So wrs can be used to test if a MB family is consistent with a DAG, it would be good if we can also find out how many consistent DAGs or essential graphs are there for this MB family.
5. also it would be good if we can explore wrs into details, such as what dag nodes become simplicial nodes in wrs recursion, and if no edges need to be deleted from a simplicial node's neighbours then what's this simplicial node?
6. maybe there is a path from s.t. every step is a moral graph of a dag, perhaps can be proved by delete an edge from a dag.

Questions: If a graph F is known to be wrs, does is help to decide if a subgraph/-supergraph different by one edge from F is wrs or not?

Answer: Probably not. If it is, then we know a base case, any graph can be reached from this base case, hence any graph can be efficiently tested.

4 Notations

4.1 Notations

References

[Diestel, 1987] R. Diestel. Simplicial decompositions of graphsfi!?some uniqueness results. *Journal of Combinatorial Theory, Series B*, 42(2):133–145, 1987.

Table 2: Notations

F	a undirected graph
G	a DAG
X	a set of random variables (nodes)
X_i	a random variable (or node) in X
X_{-i}	$X \setminus X_i$
$X_{-[1,\ldots,i]}$	$X \setminus \{X_1, \ldots, X_i\}$
B_i^G	the Markov blanket of a variable X_i in G
B_X^G	$\{B_i \mid \forall X_i \in X\}$

[Dor and Tarsi, 1992] D. Dor and M. Tarsi. A simple algorithm to construct a consistent extension of a partially oriented graph. *Technicial Report R-185, Cognitive Systems Laboratory, UCLA*, 1992.

[Gillispie and Perlman, 2001] S. B. Gillispie and M. D. Perlman. Enumerating Markov equivalence classes of acyclic digraph models. In *Proceedings of the 17th conference on Uncertainty in Artificial Intelligence*, pages 171–177. Morgan Kaufmann Publishers Inc., 2001.

[Halin, 1964] R. Halin. Über simpliziale Zerfällungen beliebiger (endlicher oder unendlicher) Graphen. *Mathematische Annalen*, 156(3):216–225, 1964.

[Verma and Pearl, 1992] T. Verma and J. Pearl. An algorithm for deciding if a set of observed independencies has a causal explanation. In *Uncertainty in Artificial Intelligence, 1992*, pages 323–330. Elsevier, 1992.

[Wagner, 1937] K. Wagner. Über eine Eigenschaft der ebenen Komplexe. *Mathematische Annalen*, 114(1):570–590, 1937.

[Wormald, 1985] N. C. Wormald. Counting labelled chordal graphs. *Graphs and combinatorics*, 1(1):193–200, 1985.