

## 6<sup>th</sup> Lecture : Structural aspects of chordal graphs and their relationships with perfect phylogeny MPRI 2013–2014

Michel Habib [habib@liafa.univ-Paris-Diderot.fr](mailto:habib@liafa.univ-Paris-Diderot.fr)  
<http://www.liafa.univ-Paris-Diderot.fr/~habib>

Sophie Germain, 29 octobre 2013

## Schedule

Introduction

Representation of chordal graphs

More structural insights of chordal graphs

Properties of reduced clique graphs

### Definition

A graph is chordal iff it has no chordless cycle of length  $\geq 4$  or equivalently it has no induced cycle of length  $\geq 4$ .

**Chordal graphs are hereditary**

Interval graphs are chordal

Applications :

- ▶ Perfect phylogeny. Many NP-complete problems for general graphs are polynomial for chordal graphs.
- ▶ Graph theory. Treewidth (resp. pathwidth) are very important graph parameters that measure distance from a chordal graph (resp. interval graph).

## Lexicographic Breadth First Search (LexBFS)

**Data:** a graph  $G = (V, E)$  and a start vertex  $s$

**Result:** an ordering  $\sigma$  of  $V$

Assign the label  $\emptyset$  to all vertices

$label(s) \leftarrow \{n\}$

**for**  $i \leftarrow n$  **à**  $1$  **do**

    Pick an unnumbered vertex  $v$  **with lexicographically largest label**

$\sigma(i) \leftarrow v$

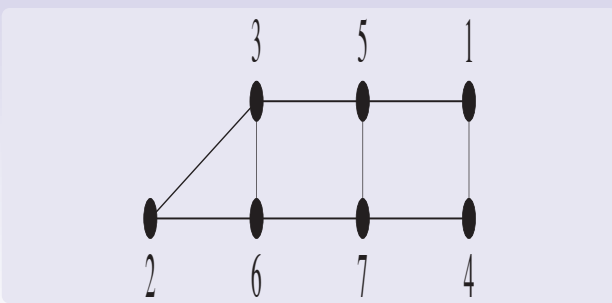
**foreach** unnumbered vertex  $w$  adjacent to  $v$  **do**

$label(w) \leftarrow label(w). \{i\}$

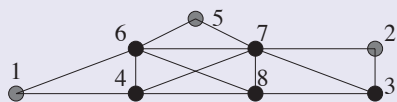
**end**

**end**

It is just a Breadth First Search (BFS for short) with a strange tie-break rule!



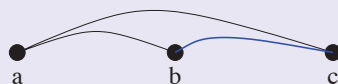
### Chordal graph



A vertex is simplicial if its neighbourhood is a clique.

### Simplicial elimination scheme

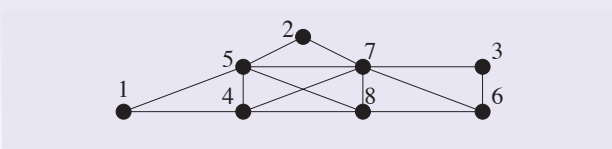
$\sigma = [x_1 \dots x_i \dots x_n]$  is a simplicial elimination scheme if  $x_i$  is simplicial in the subgraph  $G_i = G[\{x_j \dots x_n\}]$



## The reference for a graph algorithm theorem

### LexBFS Characterization [Rose, Tarjan et Lueker 1976]

A graph is chordal  $G$  iff every LexBFS ordering of  $G$  provides a simplicial elimination scheme.



How can we prove such a theorem ?

1. A direct proof, finding the invariants ?
2. Find some structure of chordal graphs
3. Understand how LexBFS explores a chordal graph
4. We will consider the 3 viewpoints.

Playing with the elimination scheme

Easy Exercises :

1. Find a minimum Coloring (resp. a clique of maximum size) of a chordal graph in  $O(|V| + |E|)$ .  
 Consequences : **chordal graphs are perfect.**  
 At most  $|V| - 1$  maximal cliques (best upper bound, since stars have exactly  $|V| - 1$  maximal cliques).
2. Find a minimum Coloring (resp. a clique of maximum size) of an interval graph in  $O(|V|)$  using the interval representation.

Greedy colorings

Definitions

Clique number  $\omega(G)$  = maximum size of a clique in  $G$   
 Chromatic number  $\chi(G)$  = minimum coloring of  $G$ .  
 $\forall G, \chi(G) \geq \omega(G)$

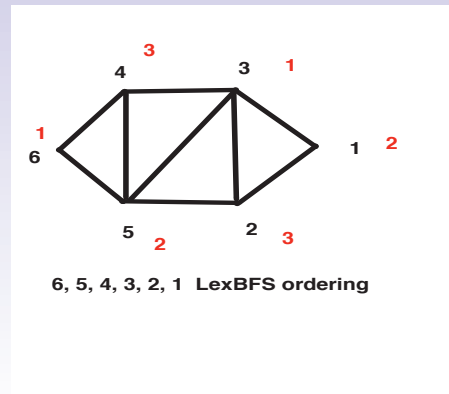
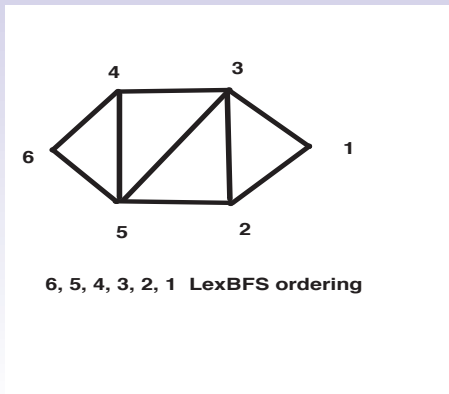
Greedy colorings

Color with integers from  $[1, k]$   
 Following a vertex ordering, process successively the vertices using the greedy rule :

Take the minimum color not already in the neighbourhood

Chordal graphs

Apply LexBFS from  $n$  downto 1  
 Use the ordering  $n$  downto 1 for the greedy coloring.  
 Let  $k = \omega(G)$ .  
 Since every added vertex  $x$  is simplicial and  $|N(x)| \leq k - 1$ , it exists at least one missing color in its neighbourhood of the already colored subgraph.  
 The value  $k$  is reached for the last vertex belonging to each maximum clique of  $G$ .



Perfect Graphs

$G$  such that for every induced subgraphs  $H \subseteq G$   
 $\omega(G) = \chi(G)$

Consequences

Therefore  $\omega(G) = \chi(G)$  for chordal graphs.  
 Since being chordal graphs is an hereditary property, chordal graphs are perfect.

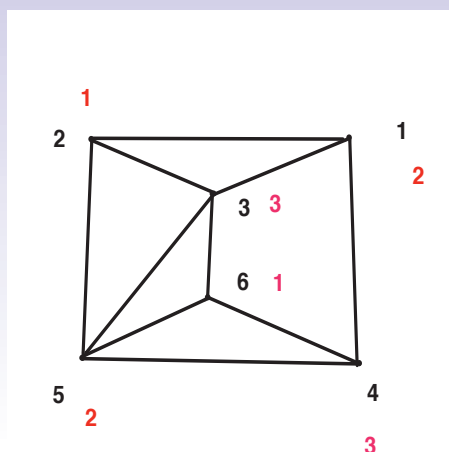
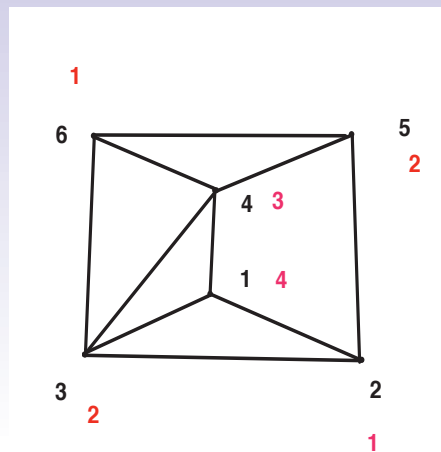
Perfectly orderable graphs

Although  $\omega(G)$  and  $\chi(G)$  can be computed in polynomial time for perfect graphs using the ellipsoid method, greedy coloring does not work for all perfect graphs.

A graph  $G$  is said to be **perfectly orderable** if there exists an ordering  $\pi$  of the vertices of  $G$ , such that any induced subgraph is optimally colored by the greedy algorithm using the subsequence of  $\pi$  induced by the vertices of the subgraph.  
 Chordal graphs are perfectly orderable.

## For which graphs the greedy coloring works ?

Bad news :  
 NP-complete to recognize perfectly orderable graphs.  
 Greedy coloring can be far from the optimum, even for subclasses of perfect graphs.



The study of the relationships between  $\omega(G)$  and  $\chi(G)$  is fundamental for algorithmic graph theory.

- 1930 Wagner's conjecture and treewidth
- 1950 Shannon Problem and Perfect graphs and semi-definite programming

## About Representations

- ▶ Interval graphs are chordal graphs
- ▶ How can we represent chordal graphs?
- ▶ As an intersection of some family?
- ▶ This family must generalize intervals on a line

## Fundamental objects to play with

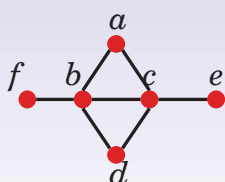
### Maximal Cliques

under inclusion

### Minimal Separators

A subset of vertices  $S$  is a **minimal separator** if  $G$  if there exist  $a, b \in G$  with  $ab \notin G$ , such that  $a$  and  $b$  are not connected in  $G - S$ .  
 and  $S$  is minimal for inclusion with this property .

## An example



3 minimal separators  $\{b\}$  for  $f$  and  $a$ ,  $\{c\}$  for  $a$  and  $e$  and  $\{b, c\}$  for  $a$  and  $d$ .

If  $G = (V, E)$  is connected then for every  $a, b \in V$  such that  $ab \notin E$   
 then there exists at least one minimal separator.  
 But there could be an **exponential number of minimal separators**.

## VIN : Maximal Clique trees

A maximal clique tree (clique tree for short) is a tree  $T$  that satisfies the following three conditions :

- ▶ Vertices of  $T$  are associated with the maximal cliques of  $G$
- ▶ Edges of  $T$  correspond to minimal separators.
- ▶ For any vertex  $x \in G$ , the cliques containing  $x$  yield a subtree of  $T$ .

## Helly Property

### Definition

A subset family  $\{T_i\}_{i \in I}$  satisfies Helly property if  $\forall J \subseteq I$  et  $\forall i, j \in J$   $T_i \cap T_j \neq \emptyset$  implies  $\bigcap_{i \in I} T_i \neq \emptyset$

### Exercise

Subtrees in a tree satisfy Helly property.

### Démonstration.

Suppose not. Consider a family of subtrees that pairwise intersect. For each vertex  $x$  of the tree  $T$ , if  $x$  belongs to every subtree of the family, it contradicts the hypothesis. Therefore at least one subtree does not contain  $x$ . If the subtrees belongs to two different components of  $T-x$  this would contradict the pairwise intersection of the subtrees. Therefore all the subtrees are in exactly one component of  $T-x$  (N.B. some subtrees may contain  $x$ ). Direct exactly one edge of  $T$  from  $x$  to this component. This yields a directed graph  $G$ , which has exactly  $n$  vertices and  $n$  directed edges. Since  $T$  is a tree, it contains no cycle, therefore it must exist a pair of symmetric edges in  $G$ , which contradicts the pairwise intersection. □

## Back to chordal graphs

### Chordal graph recognition

1. Apply a LexBFS on  $G$   $O(n+m)$
2. Check if the reverse ordering is a simplicial elimination scheme  $O(n+m)$
3. In case of failure, exhibit a certificate : i.e. a cycle of length  $\geq 4$ , without a chord.  $O(n)$

## Subtrees in a tree

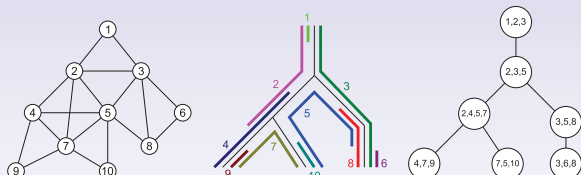
Using results of Dirac 1961, Fulkerson, Gross 1965, Buneman 1974, Gavril 1974 and Rose, Tarjan and Lueker 1976 :

For a connected graph, the following statements are equivalent and characterize chordal graphs :

- (i)  $G$  has a simplicial elimination scheme
- (ii) Every minimal separator is a clique
- (iii)  $G$  admits a maximal clique tree.
- (iv)  $G$  is the intersection graph of subtrees in a tree.
- (v) Any MNS (LexBFS, LexDFS, MCS) provides a simplicial elimination scheme.

Two subtrees intersect iff they have at least one vertex in common.  
 By no way, these representations can be uniquely defined !

## An example



## Proof of the chordal characterization theorem

- ▶ Clearly (iii) implies (iv)
- ▶ For the converse, each vertex of the tree corresponds to a clique in  $G$ .  
 But the tree has to be pruned of all its unnecessary nodes, until in each node some subtree starts or ends. Then nodes correspond to maximal cliques.
- ▶ We need now to relate these new conditions to chordal graphs.
  - (iii) implies (i) since a maximal clique tree yields a simplicial elimination scheme.
  - (iv) implies chordal since a cycle without a chord generates a cycle in the tree.
  - (iv) implies (ii) since each edge of the tree corresponds to a minimal separator which is a clique

from (i) to (iv)

**Démonstration.**

By induction on the number of vertices. Let  $x$  be a simplicial vertex of  $G$ .  
 By induction  $G - x$  can be represented with a family of subtrees on a tree  $T$ .  
 $N(x)$  is a clique and using Helly property, the subtrees corresponding to  $N(x)$  have a vertex in common  $\alpha$ .  
 To represent  $G$  we just add a pending vertex  $\beta$  adjacent to  $\alpha$ .  
 $x$  being represented by a path restricted to the vertex  $\beta$ , and we add to all the subtrees corresponding to vertices in  $N(x)$  the edge  $\alpha\beta$ . □

**Exercises**

1. Can we use efficiently this representation of chordal graphs as intersection of subtrees ?
2. Same question for path graphs ? (intersection graph of paths in a tree)

**First remark**

**Proposition**

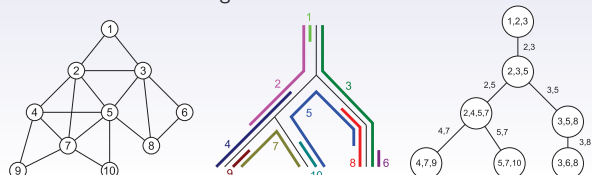
Every undirected graph can be obtained as the intersection of a subset family

**Proof**

$G = (V, E)$   
 Let us denote by  $E_x$  the set of edges adjacent to  $x$ .  
 $xy \in E$  iff  $E_x \cap E_y \neq \emptyset$   
 We could also have taken the set  $C_x$  of all maximal cliques which contains  $x$ .  
 $C_x \cap C_y \neq \emptyset$  iff  $\exists$  one maximal clique containing both  $x$  and  $y$

**Clique tree**

*clique tree* of  $G =$  a minimum size tree model of  $G$   
 for a clique tree  $T$  of  $G$  :  
 ▶ vertices of  $T$  correspond precisely to the maximal cliques of  $G$   
 ▶ for every maximal cliques  $C, C'$ , each clique on the path in  $T$  from  $C$  to  $C'$  contains  $C \cap C'$   
 ▶ for each edge  $CC'$  of  $T$ , the set  $C \cap C'$  is a *minimal separator* (an inclusion-wise minimal set separating two vertices)  
 Note : we label each edge  $CC'$  of  $T$  with the set  $C \cap C'$ .



**Playing with the representation**

**Easy Exercises :**

1. Find a minimum Coloring (resp. a clique of maximum size) of a chordal graph in  $O(n + m)$ .  
 Consequences : **chordal graphs are perfect.**  
 At most  $n$  maximal cliques.
2. Find a minimum Coloring (resp. a clique of maximum size) of an interval graph in  $O(n)$  using the interval representation.

**Which kind of representation to look for for special classes of graphs ?**

- ▶ Easy to manipulate (optimal encoding, easy algorithms for optimisation problems)
- ▶ Geometric in a wide meaning (ex : permutation graphs = intersection of segments between two lines)
- ▶ Examples : disks in the plane, circular genomes ...

Starting from a graph in some application, find its characteristic :

1. 2-intervals on a line (biology), intersection of disks (or hexagons) in the plane (radio frequency), filament graphs, trapezoid graphs ...
2. A whole book on this subject :  
 J. Spinrad, Efficient Graph Representations, Fields Institute Monographs, 2003.
3. A website on graph classes :  
<http://www.graphclasses.org/>

**Consequences of maximal clique tree**

**Theorem**  
 Every minimal separator belongs to every maximal clique tree.

**Lemma**  
 Every minimal separator is the intersection of at least 2 maximal cliques of  $G$ .

**Corollary**  
 There are at most  $n$  minimal separators.

**Theorem**

Every minimal separator belongs to every maximal clique tree.

**Lemma**

Every minimal separator is the intersection of at least 2 maximal cliques of  $G$ .

**Proof of the lemma**

**Démonstration.**

Since  $G$  is chordal, every minimal separator  $S$  is a clique. Let us consider  $G_1$  a connected component of  $G-S$ . Let  $x_1, \dots, x_k$  be the vertices of  $G_1$  having a maximal neighbourhood in  $S$ . If  $k = 1$  then  $x_1$  must be universal to  $S$ , since  $S$  is a minimal separator. Else, consider a shortest path  $\mu$  in  $G_1$  from  $x_1$  to  $x_k$ . Necessarily  $x_1$  (resp.  $x_k$ ) has a private neighbour  $z$  (resp.  $t$ ) in  $S$ . Else they would have the same maximal neighbourhood in  $S$ , and since  $S$  is a minimal separator, this neighbourhood must be  $S$ . Then the cycle  $[x_1, \mu, x_k, t, z]$  has no chord, a contradiction. Therefore  $x_1 \cup S$  is a clique, and is contained in some maximal clique  $C$  in  $G_1$ . We finish the proof by considering another connected component of  $G - S$ . □

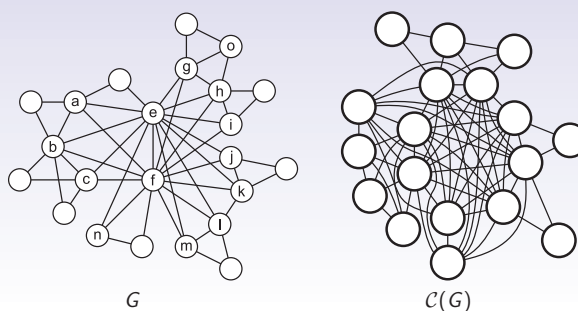
**Proof of the theorem**

**Démonstration.**

Therefore  $S = C' \cap C''$ . These two maximal cliques belong to any maximal clique tree  $T$  of  $G$ . Let us consider the unique path  $\mu$  in  $T$  joining  $C'$  to  $C''$ . All the internal maximal cliques in  $\mu$  must contain  $S$ . Suppose that all the edges of  $\mu$  are labelled with minimal separators strictly containing  $S$ , then we can construct a path in  $G$  from  $C' - S$  to  $C'' - S$  avoiding  $S$ , a contradiction. So at least one edge of  $\mu$  is labelled with  $S$ . □

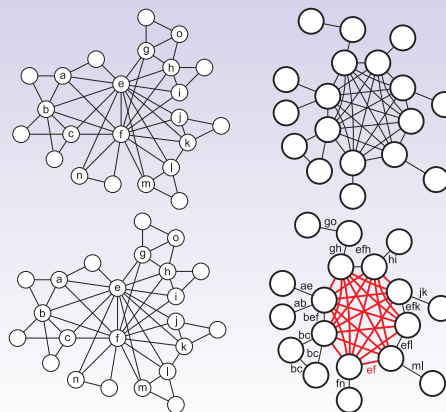
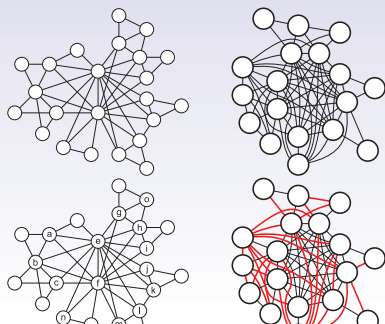
**Clique graph**

the *clique graph*  $\mathcal{C}(G)$  of  $G =$  intersection graph of maximal cliques of  $G$



**Reduced clique graph**

the *reduced clique graph*  $\mathcal{C}_r(G)$  of  $G =$  graph on maximal cliques of  $G$  where  $CC'$  is an edge of  $\mathcal{C}_r(G) \iff C \cap C'$  is a minimal separator.



**Combinatorial structure of  $\mathcal{C}_r(G)$**

**Lemma 1 : M.H and C. Paul 95**

If  $C_1, C_2, C_3$  is a cycle in  $\mathcal{C}_r(G)$ , with  $S_{12}, S_{23}$  and  $S_{31}$  be the associated minimal separators then two of these three separators are equal and included in the third.

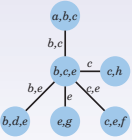
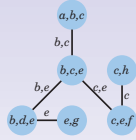
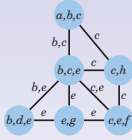
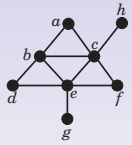
**Lemma 2 : M.H. and C. Paul 95**

Let  $C_1, C_2, C_3$  be 3 maximal cliques, if  $C_1 \cap C_2 = S_{12} \subset S_{23} = C_2 \cap C_3$  then it yields a triangle in  $\mathcal{C}_r(G)$

**Lemma 3 : Equality case**

Let  $C_1, C_2, C_3$  be 3 maximal cliques, if  $S_{12} = S_{23}$  then :

- ▶ either the  $C_1 \cap C_3 = S_{13}$  is a minimal separator
- ▶ or the edges  $C_1 C_2$  and  $C_2 C_3$  cannot belong together to a maximal clique tree of  $G$ .



**Theorem (Gavril 87, Shibata 1988, Blayr and Payton 93)**

The clique trees of  $G$  are precisely the maximum weight spanning trees of  $\mathcal{C}(G)$  where the weight of an edge  $CC'$  is defined as  $|C \cap C'|$ .

**Theorem (Galinier, Habib, Paul 1995)**

The clique trees of  $G$  are precisely the maximum weight spanning trees of  $\mathcal{C}_r(G)$  where the weight of an edge  $CC'$  is defined as  $|C \cap C'|$ .

Moreover,  $\mathcal{C}_r(G)$  is the union of all clique trees of  $G$ .

**Applications**

- ▶ All clique trees have exactly the same labels, including repetitions.

**Maximal Cardinality Search : MCS**

**Data:** a graph  $G = (V, E)$  and a start vertex  $s$

**Result:** an ordering  $\sigma$  of  $V$

Assign the label 0 to all vertices

$label(s) \leftarrow 1$

**for**  $i \leftarrow n$  **à** 1 **do**

    Pick an unnumbered vertex  $v$  with largest label

$\sigma(i) \leftarrow v$

**foreach** unnumbered vertex  $w$  adjacent to  $v$  **do**

$label(w) \leftarrow label(w) + 1$

**end**

**end**

**Maximum spanning trees**

Maximal Cardinality Search can be seen as Prim algorithm for computing a maximal spanning tree of  $\mathcal{C}_r(G)$ .