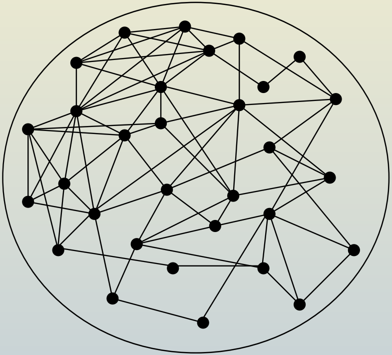
The complexity of finding and counting small subgraphs

OU Winter Combinatorics Meeting, 20th January 2016

Kitty Meeks

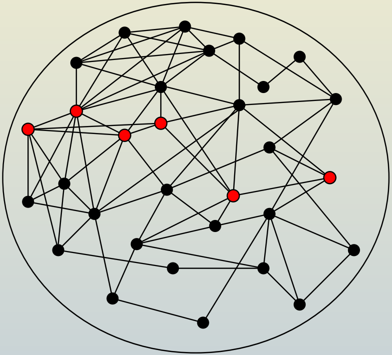
The problem

Given a graph on n vertices, we are interested in subgraphs with k vertices that have particular properties.



The problem

Given a graph on n vertices, we are interested in subgraphs with k vertices that have particular properties.

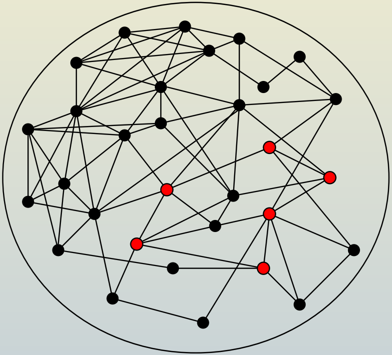


For example:

- Paths on k vertices

The problem

Given a graph on n vertices, we are interested in subgraphs with k vertices that have particular properties.

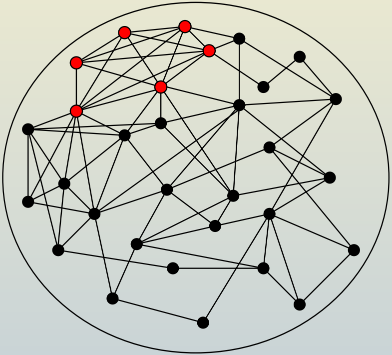


For example:

- Paths on k vertices
- Cycles on k vertices

The problem

Given a graph on n vertices, we are interested in subgraphs with k vertices that have particular properties.

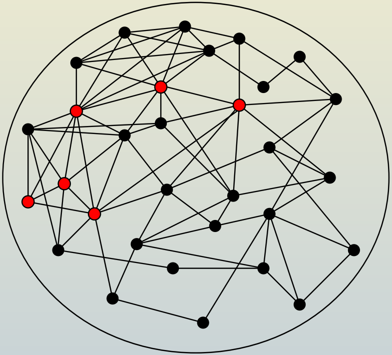


For example:

- Paths on k vertices
- Cycles on k vertices
- Cliques on k vertices

The problem

Given a graph on n vertices, we are interested in subgraphs with k vertices that have particular properties.

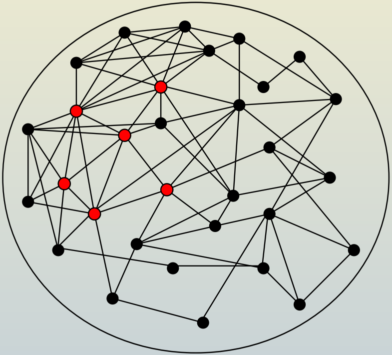


For example:

- Paths on k vertices
- Cycles on k vertices
- Cliques on k vertices
- Connected k-vertex induced subgraphs

The problem

Given a graph on n vertices, we are interested in subgraphs with k vertices that have particular properties.



For example:

- Paths on k vertices
- Cycles on k vertices
- Cliques on k vertices
- Connected k-vertex induced subgraphs
- k-vertex induced subgraphs with an even number of edges

DECISION

Is there a witness?

Deciding, counting and enumerating

DECISION
Is there a witness?

APPROX COUNTING
Approximately how many witnesses?

DECISION
Is there a witness?

APPROX COUNTING
Approximately how many witnesses?

EXACT COUNTING
Exactly how many witnesses?

Deciding, counting and enumerating

DECISION
Is there a witness?

EXTRACTION
Identify a single witness

APPROX COUNTING
Approximately how many witnesses?

EXACT COUNTING
Exactly how many witnesses?

Deciding, counting and enumerating

DECISION
Is there a witness?

EXTRACTION
Identify a single witness

APPROX COUNTING
Approximately how many witnesses?

UNIFORM SAMPLING
Pick a single witness uniformly at random

EXACT COUNTING
Exactly how many witnesses?

Deciding, counting and enumerating

DECISION
Is there a witness?

EXTRACTION
Identify a single witness

APPROX COUNTING
Approximately how many witnesses?

UNIFORM SAMPLING
Pick a single witness uniformly at random

EXACT COUNTING
Exactly how many witnesses?

ENUMERATION
List all witnesses

The values of n and k

Consider the k-cycle problem:

- if $k = 3$ then we are interested in triangles
- if $k = n$ then we are interested in Hamilton Cycles

The values of n and k

Consider the k-cycle problem:

- if $k = 3$ then we are interested in triangles
- if $k = n$ then we are interested in Hamilton Cycles

We are interested in what happens as n and k both tend to infinity, independently, with $k << n$.

The values of n and k

Consider the k-cycle problem:

- if $k = 3$ then we are interested in triangles
- if $k = n$ then we are interested in Hamilton Cycles

We are interested in what happens as n and k both tend to infinity, independently, with $k << n$.

- We can consider all possible k-vertex subgraphs in time $O(n^k)$.

The values of n and k

Consider the k-cycle problem:

- if $k = 3$ then we are interested in triangles
- if $k = n$ then we are interested in Hamilton Cycles

We are interested in what happens as n and k both tend to infinity, independently, with $k << n$.

- We can consider all possible k-vertex subgraphs in time $O(n^k)$.
- We would like to be able to answer questions about k-vertex subgraphs in time $f(k) \cdot n^{O(1)}$.

Deciding, counting and enumerating

DECISION
Is there a witness?

EXTRACTION
Identify a single witness

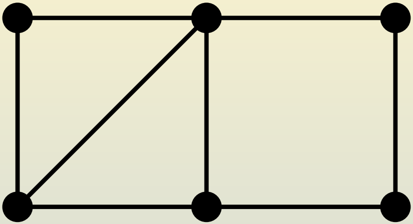
APPROX COUNTING
Approximately how many witnesses?

UNIFORM SAMPLING
Pick a single witness uniformly at random

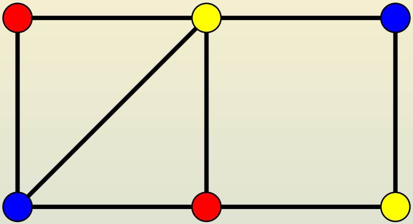
EXACT COUNTING
Exactly how many witnesses?

ENUMERATION
List all witnesses

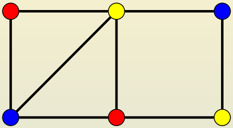
Hardness, using poset lattices

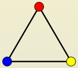


Hardness, using poset lattices

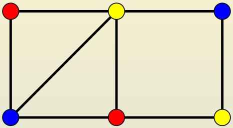


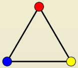
Hardness, using poset lattices





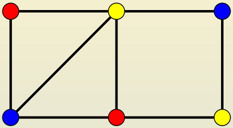
Hardness, using poset lattices

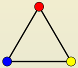




$$\text{mask}\left\{\overbrace{\begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}}^{\text{underlying structure}}\right. \cdot \begin{pmatrix} \# \\ \# \\ \# \\ \# \\ \# \\ \# \\ \# \\ \# \end{pmatrix} = \begin{pmatrix} \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}$$

Hardness, using poset lattices

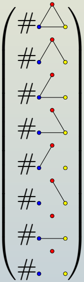




underlying structure

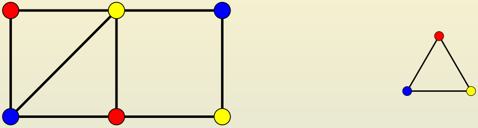
mask

$$\begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \cdot$$



$$= \begin{pmatrix} 4 \\ 4 \\ 2 \\ 6 \\ 4 \\ 2 \\ 2 \end{pmatrix}$$

Hardness, using poset lattices



$$
\text{mask}\left\{
\overbrace{\begin{pmatrix}
0 & 1 & 1 & 1 & 0 & 0 & 0 & 1\\
1 & 1 & 0 & 0 & 0 & 0 & 1 & 1\\
1 & 0 & 1 & 0 & 0 & 1 & 0 & 1\\
1 & 0 & 0 & 1 & 1 & 0 & 0 & 1\\
0 & 0 & 0 & 1 & 0 & 1 & 1 & 1\\
0 & 0 & 1 & 0 & 1 & 0 & 1 & 1\\
0 & 1 & 0 & 0 & 1 & 1 & 0 & 1\\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{pmatrix}}^{\text{underlying structure}}\right.
\cdot
\begin{pmatrix}
\#\\ \#\\ \#\\ \#\\ \#\\ \#\\ \#\\ \#
\end{pmatrix}
=
\begin{pmatrix}
4\\ 4\\ 2\\ 6\\ 4\\ 2\\ 2
\end{pmatrix}
$$

Method used by: Jerrum & M. ('14, '16); matrix inversion more generally used by Flum & Grohe ('04), Bläser & Curticapean ('12), Curticapean ('13).

Tractability, using Ramsey

- Does G contain a set of k vertices that induces either a clique or independent set?

Tractability, using Ramsey

- Does G contain a set of k vertices that induces either a clique or independent set?

- How many such subsets does G contain?

- Does G contain a set of k vertices that induces either a clique or independent set?
- How many such subsets does G contain?

Proposition

Let $G = (V, E)$ be an n-vertex graph, where $n \geq 2^{2k}$. Then the number of k-vertex subsets $U \subset V$ such that U induces either a clique or independent set in G is at least

$$\frac{(2^{2k} - k)!}{(2^{2k})!} \frac{n!}{(n-k)!}.$$

Tractability, using Ramsey

- Does G contain a set of k vertices that induces either a clique or independent set?
- How many such subsets does G contain?

Proposition

Let $G = (V, E)$ be an n-vertex graph, where $n \geq 2^{2k}$. Then the number of k-vertex subsets $U \subset V$ such that U induces either a clique or independent set in G is at least

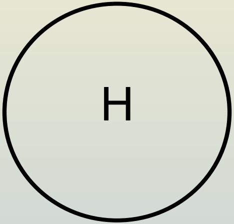
$$\frac{(2^{2k} - k)!}{(2^{2k})!} \frac{n!}{(n-k)!}.$$

Method used by: Arvind & Raman ('02), Khot & Raman ('02), Jerrum & M. ('16).

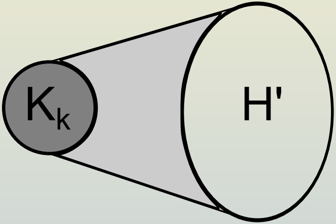
Suppose H has 2^{2k} vertices. If we can decide whether G contains an **induced** copy of H, we can decide whether G contains a k-clique.

Hardness, using Ramsey

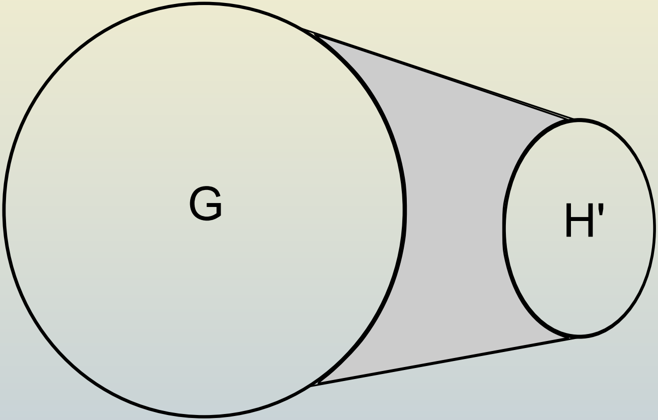
Suppose H has 2^{2k} vertices. If we can decide whether G contains an **induced** copy of H, we can decide whether G contains a k-clique.



Suppose H has 2^{2k} vertices. If we can decide whether G contains an **induced** copy of H, we can decide whether G contains a k-clique.

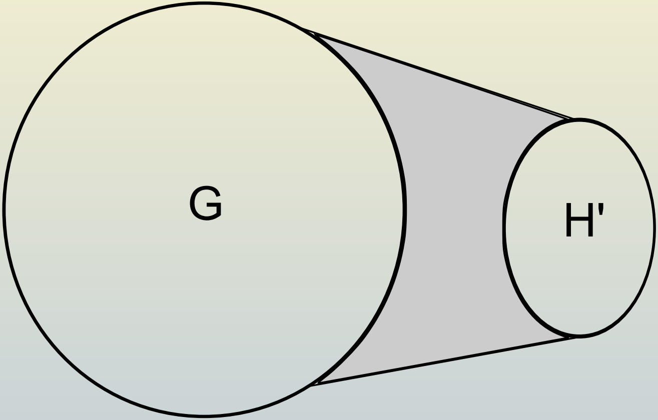


Suppose H has 2^{2k} vertices. If we can decide whether G contains an **induced** copy of H, we can decide whether G contains a k-clique.



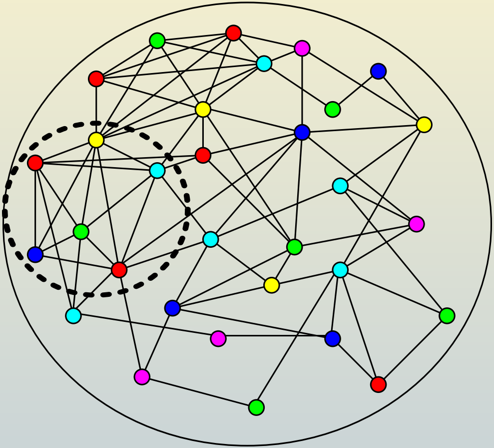
Hardness, using Ramsey

Suppose H has 2^{2k} vertices. If we can decide whether G contains an **induced** copy of H, we can decide whether G contains a k-clique.

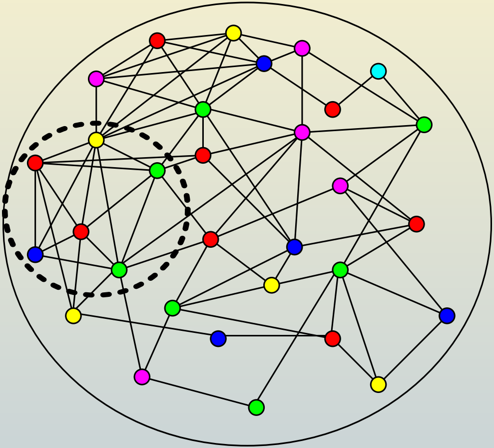


Method used by: Chen, Thurley & Weyer ('08), Khot & Raman ('02), Curticapean & Marx ('14), Jerrum & M. ('15).

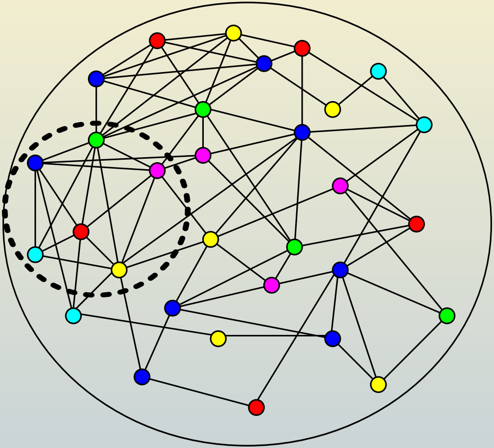
Tractability, using colour-coding



Tractability, using colour-coding



Tractability, using colour-coding

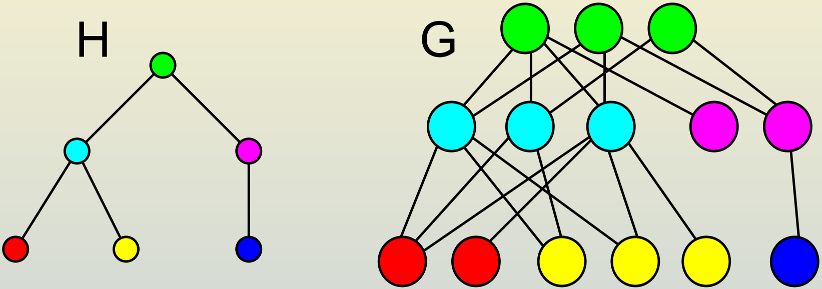


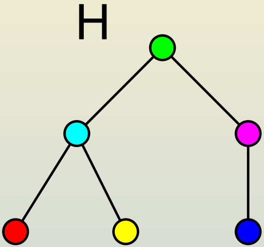
Tractability, using colour-coding

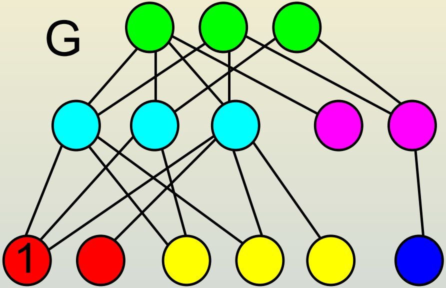
Theorem (Alon, Yuster and Zwick 1995)

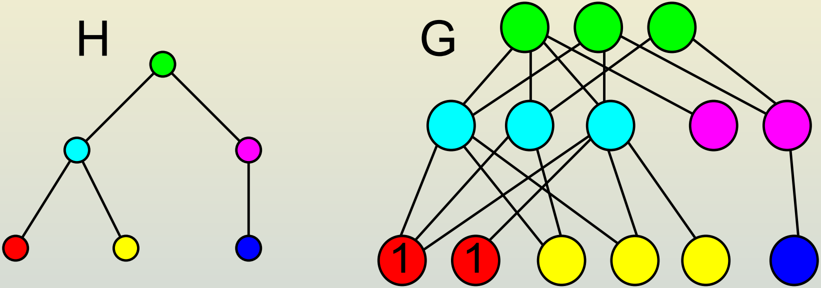
For all $n, k \in \mathbb{N}$, there is a k-perfect family $\mathcal{F}_{n,k}$ of hash functions from $[n]$ to $[k]$ of cardinality $2^{O(k)} \cdot \log n$, and (given n and k) the family $\mathcal{F}_{n,k}$ can be computed in time $2^{O(k)} \cdot n \cdot \log^2 n$.

Tractability, using bounded treewidth

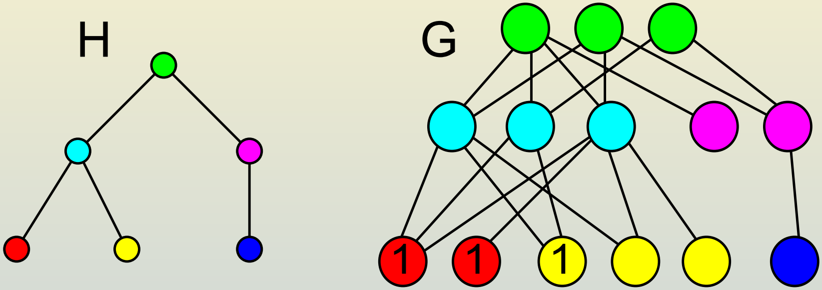


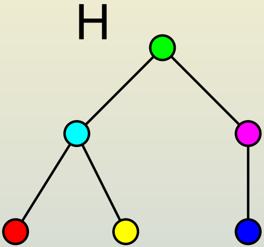

H

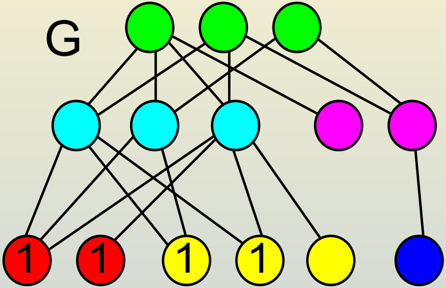

G
1


H
G
1
1

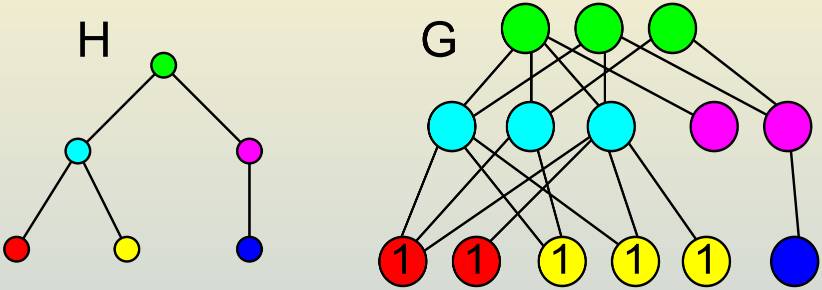
Tractability, using bounded treewidth



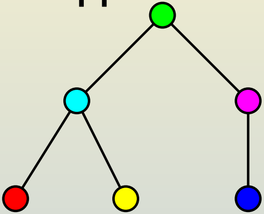

H


G
1
1
1
1

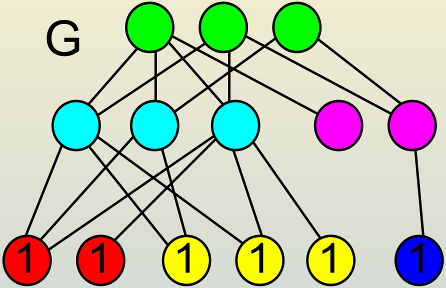
Tractability, using bounded treewidth



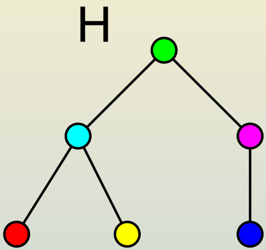
H

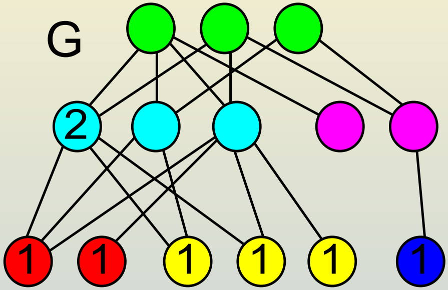


G

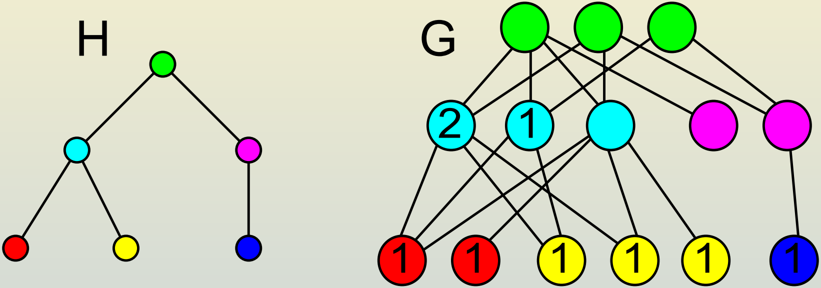


Tractability, using bounded treewidth

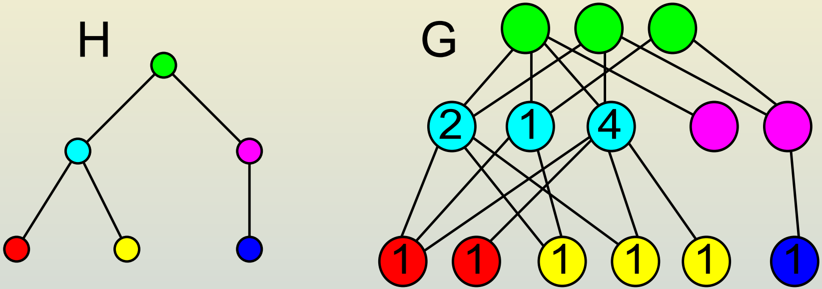




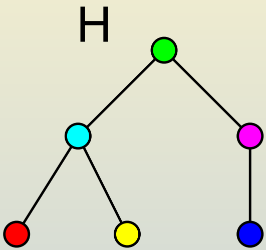
Tractability, using bounded treewidth

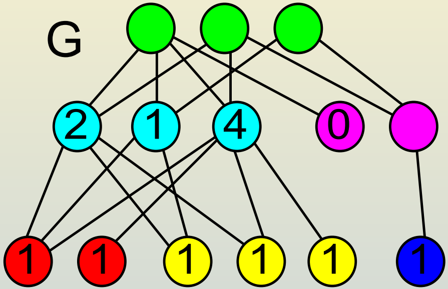


Tractability, using bounded treewidth

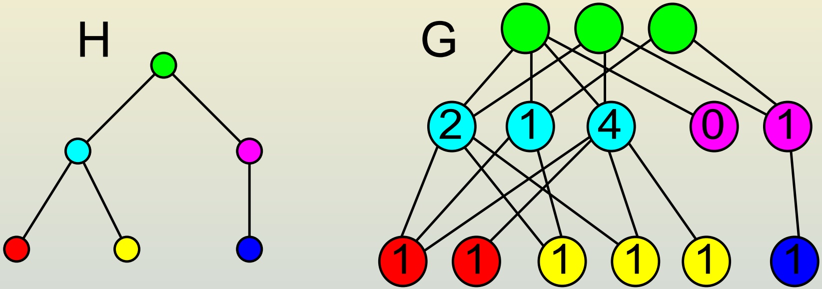


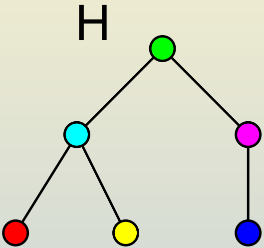
Tractability, using bounded treewidth

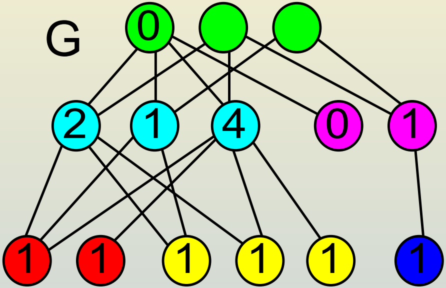




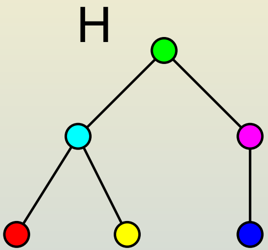
Tractability, using bounded treewidth

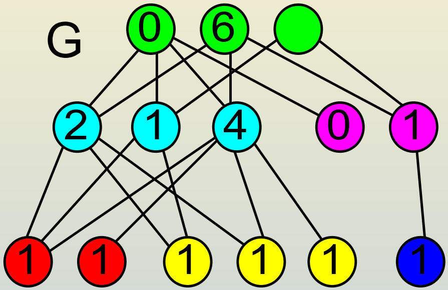



H

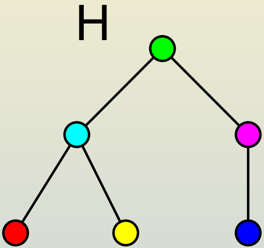

G
0
2
1
4
0
1
1
1
1
1
1
1

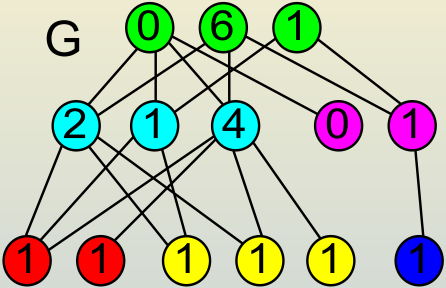
Tractability, using bounded treewidth

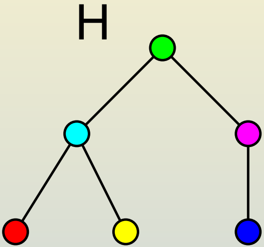


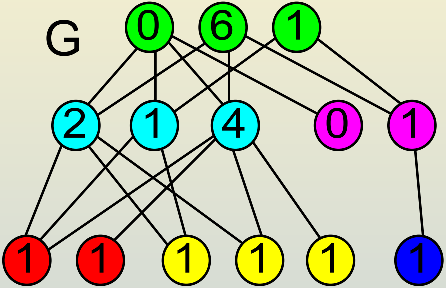


Tractability, using bounded treewidth









Method used by: Alon, Yuster & Zwick (1995), Arvind & Raman (2002), Jerrum & M. (2015).

Hardness, using unbounded treewidth

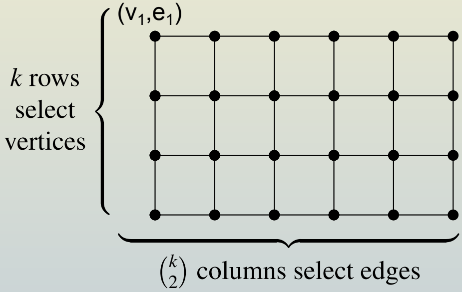
Theorem (Excluded Grid Theorem, Robertson and Seymour 1986)

There is a computable function $w : N \to N$ *such that the* $(k \times k)$ *grid is a minor of every graph of treewidth at least* $w(k)$.

Hardness, using unbounded treewidth

Theorem (Excluded Grid Theorem, Robertson and Seymour 1986)

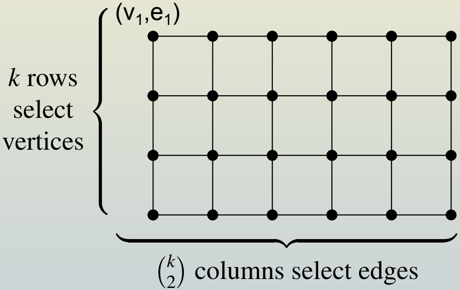
There is a computable function $w : N \to N$ such that the $(k \times k)$ grid is a minor of every graph of treewidth at least $w(k)$.



Hardness, using unbounded treewidth

Theorem (Excluded Grid Theorem, Robertson and Seymour 1986)

There is a computable function $w : N \to N$ such that the $(k \times k)$ grid is a minor of every graph of treewidth at least $w(k)$.



Method used by: Grohe, Schwentick & Segoufin (2001), Dalmau & Jonsson (2004), Grohe (2007), Färnqvist & Jonnson (2007), Chen, Thurley & Weyer (2008), Chen & Müller (2008), M. (2016).

Are there examples of problems of this form:

Open problems

Are there examples of problems of this form:

- for which exact counting is tractable? (Other than the “trivial” examples...)

Open problems

Are there examples of problems of this form:

- for which exact counting is tractable? (Other than the "trivial" examples...)
- for which decision is tractable but approximate counting is hard?

Open problems

Are there examples of problems of this form:

- for which exact counting is tractable? (Other than the “trivial” examples...)
- for which decision is tractable but approximate counting is hard?

Thank you