An FPT algorithm for counting subgraphs

Jess Ryan

April 2019

Outline

1. Introduction
   - Definitions
   - Statement of problem
   - Motivation for problem

What is the subgraph counting problem?

Problem Statement

How many (unlabelled) copies of the graph H are contained in the graph G?

We call the graph G the *host* graph and H the *pattern* graph.

Graph Theory Definitions

Almost bounded degree graph

A graph G has *almost bounded degree* k if G contains at most k vertices with degree greater than k.

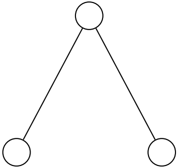
What is the subgraph counting problem?

Problem Statement

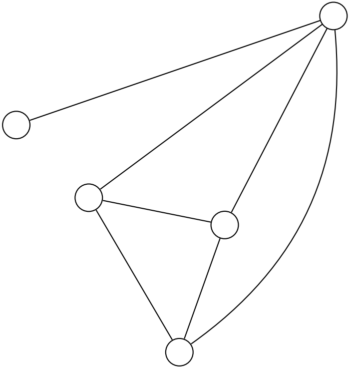
How many (unlabelled) copies of the graph H are contained in the graph G?

We call the graph G the *host* graph and H the *pattern* graph.

H

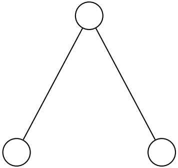


G

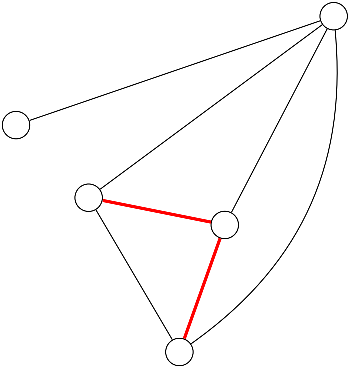


Count $=0$

H

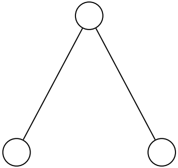


G

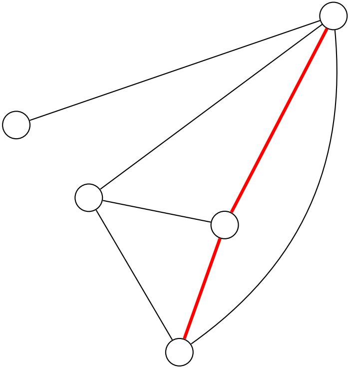


Count $=0+1$

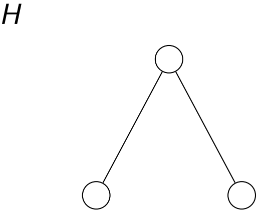
H

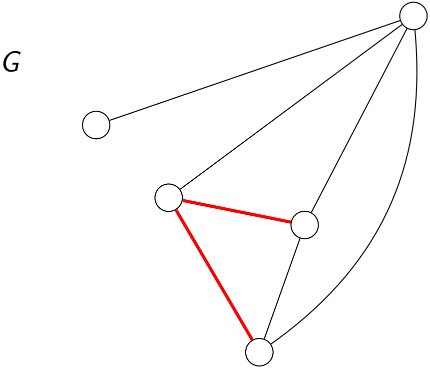


G



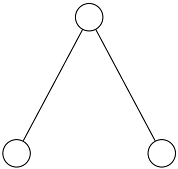
Count $= 1 + 1$





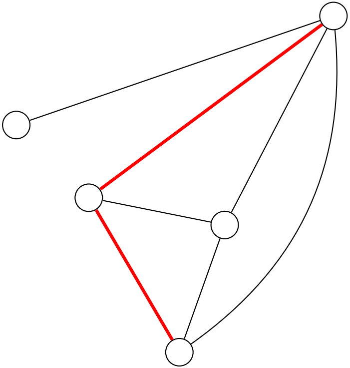
Count $= 2 + 1$

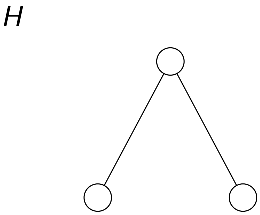
H

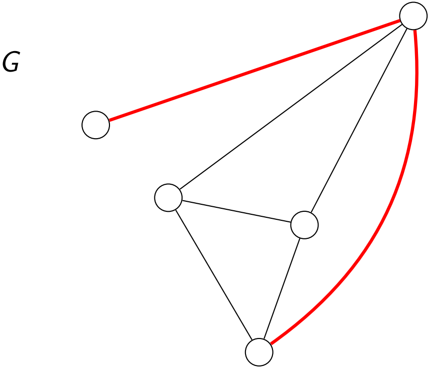


Count $=3+1$

G

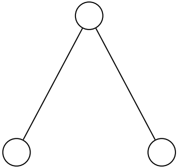




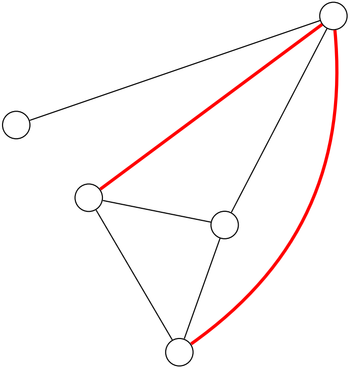


Count $=4+1$

H

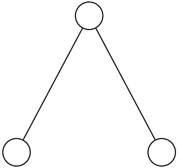


G

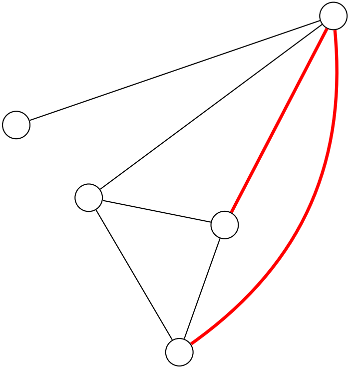


Count $=5+1$

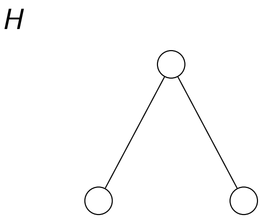
H

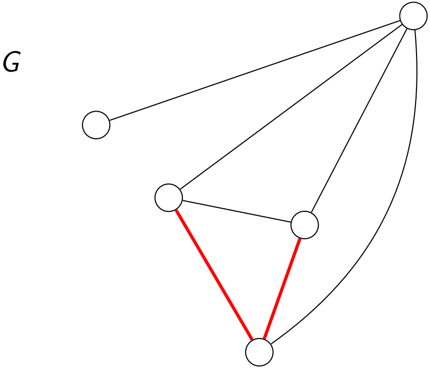


G



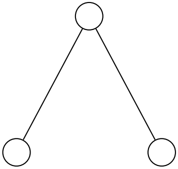
Count $=6{+}1$



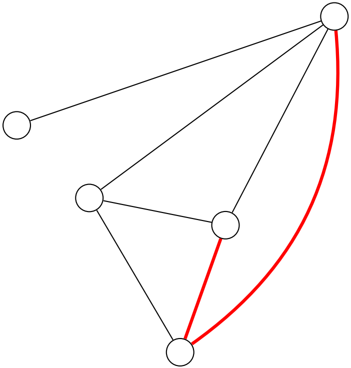


Count $= 7 + 1$

H

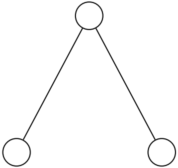


G

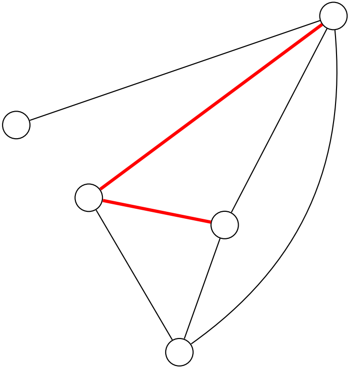


Count $= 8 + 1$

H

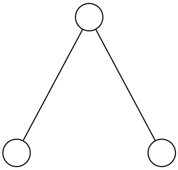


G

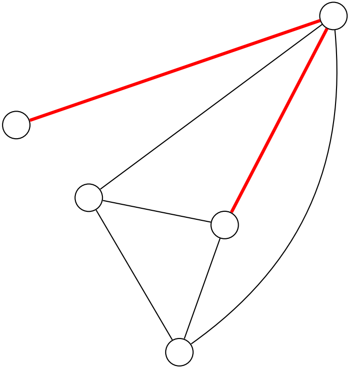


Count $=9+1$

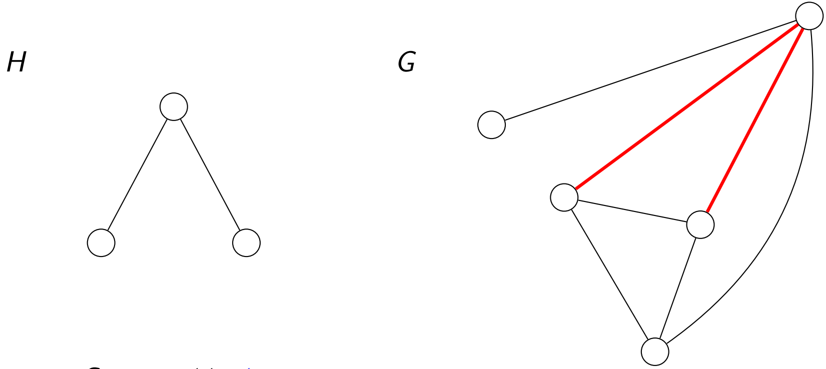
H



G

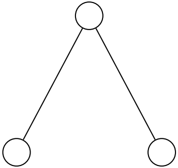


Count $=10+1$



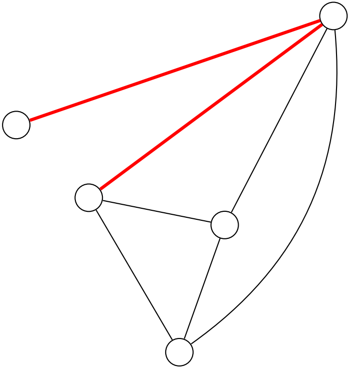
Count $=11+1$

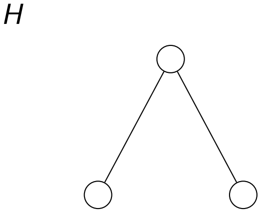
H

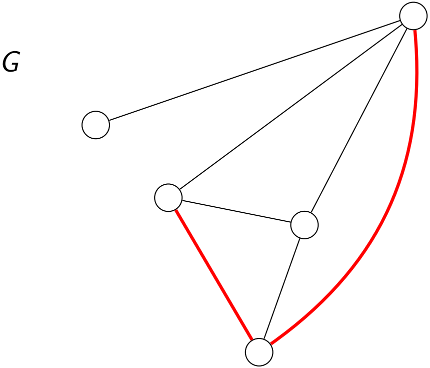


Count $=12+1$

G

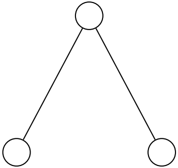




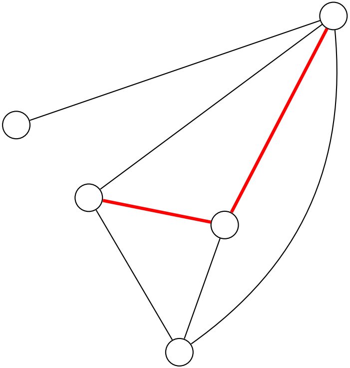


Count $=13+1$

H



G



Count $=14+1=15$

Why do we care?

- Generalisation of subgraph isomorphism

Why do we care?

- Generalisation of subgraph isomorphism
- Network analysis

Algorithmic complexity

NP-hard

The class *NP-hard* contains all problems for which no polynomial time algorithm has been found. Unless $P = NP$, which we do not expect is true, no NP-hard problem can be solved in polynomial time.

Algorithmic complexity

NP-hard

The class *NP-hard* contains all problems for which no polynomial time algorithm has been found. Unless $P = NP$, which we do not expect is true, no NP-hard problem can be solved in polynomial time.

NP-complete

A problem is in the class *NP-complete* if it is NP-hard and we can verify a solution to this problem in polynomial time.

NP-hard

The class *NP-hard* contains all problems for which no polynomial time algorithm has been found. Unless $P = NP$, which we do not expect is true, no NP-hard problem can be solved in polynomial time.

NP-complete

A problem is in the class *NP-complete* if it is NP-hard and we can verify a solution to this problem in polynomial time.

Fixed-parameter tractable (FPT)

A problem belongs to the class FPT if it can be solved in polynomial time for small, fixed values of some parameter of the problem other than its input size.

Subgraph counting is hard!

- Subgraph isomorphism is NP-complete $\rightarrow$ subgraph counting is NP-complete

Subgraph counting is hard!

- Subgraph isomorphism is NP-complete $\rightarrow$ subgraph counting is NP-complete
- Assuming Exponential Time Hypothesis $\rightarrow$, subgraph counting is not in FPT in general

Subgraph counting is hard!

- Subgraph isomorphism is NP-complete $\rightarrow$ subgraph counting is NP-complete
- Assuming Exponential Time Hypothesis $\rightarrow$, subgraph counting is not in FPT in general
- (Enright and Meeks) Subgraph counting is in FPT when the host graph has *almost bounded degree*

My PhD mini-project

Project objective:

Design and implement an FPT algorithm for subgraph counting in host graphs with almost bounded degree k.

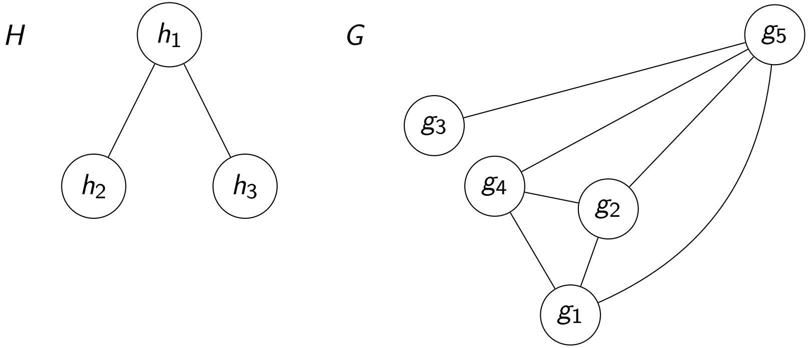
Recall:

A graph G has *almost bounded degree* k if G contains at most k vertices of degree greater than k.

Algorithm: General Idea

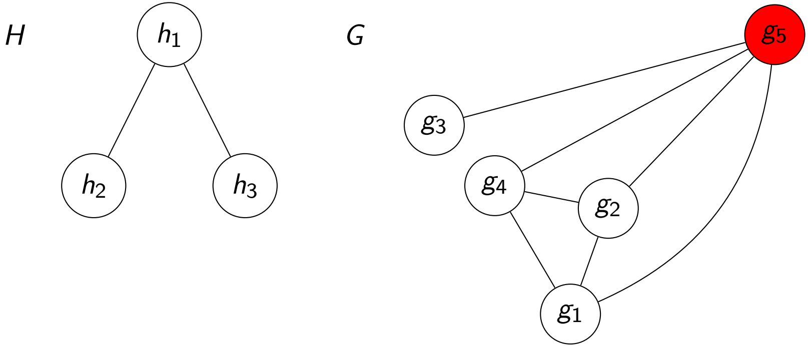
- Consider each way to assign subset H to high degree vertices in G
- For each feasible assignment, count ways to assign the rest of H to the bounded degree part of G
- Sum up the counts to obtain number of labelled copies of H in G
- Divide by the number of copies of H in H to obtain number of *unlabelled* copies of H in G

Example



$h_1 : g_1, g_2, g_3, g_4, g_5$
$h_2 : g_1, g_2, g_3, g_4, g_5$
$h_3 : g_1, g_2, g_3, g_4, g_5$

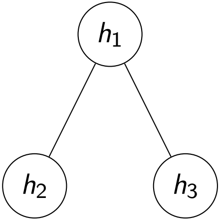
Example



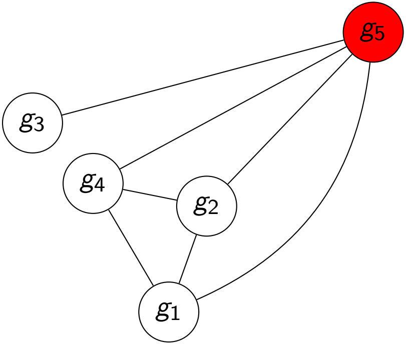
$h_1 : g_1, g_2, g_3, g_4, g_5$
$h_2 : g_1, g_2, g_3, g_4, g_5$
$h_3 : g_1, g_2, g_3, g_4, g_5$

Example

H



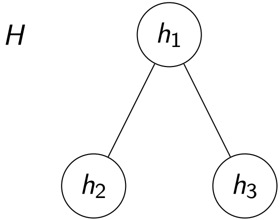
G

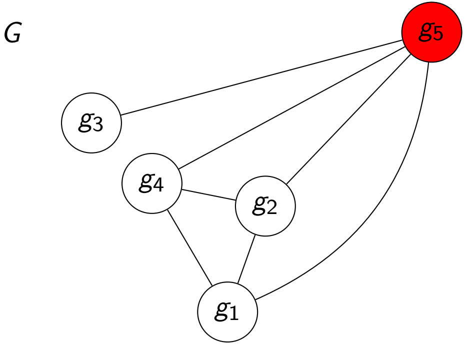


Subset of $V(H)$: $\emptyset$
Count $= 0$

$h_1: g_1, g_2, g_3, g_4, \cancel{g_5}$
$h_2: g_1, g_2, g_3, g_4, \cancel{g_5}$
$h_3: g_1, g_2, g_3, g_4, \cancel{g_5}$

Example



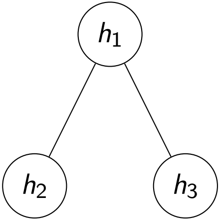


Subset of $V(H)$: $\emptyset$
Count $= 0$

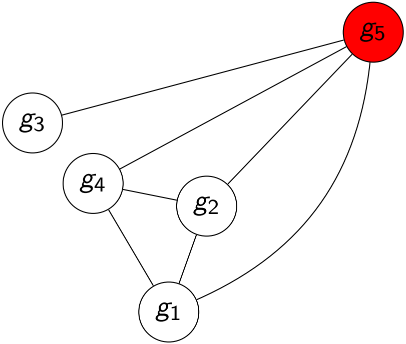
$h_1 \rightarrow g_1$
h_2 : $\cancel{g_1}, g_2, \cancel{g_3}, g_4$
h_3 : $\cancel{g_1}, g_2, \cancel{g_3}, g_4$

Example

H



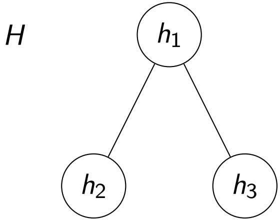
G

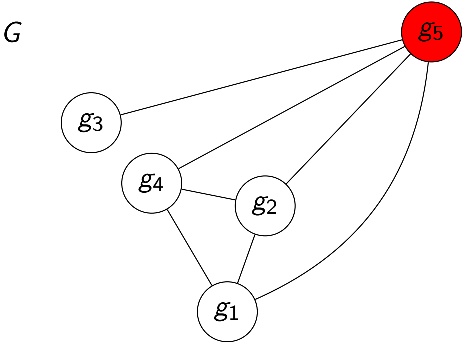


Subset of $V(H)$: $\emptyset$
Count $= 0$

$h_1 \rightarrow g_1$
$h_2 \rightarrow g_2$
$h_3 : \cancel{g_2}, g_4$

Example

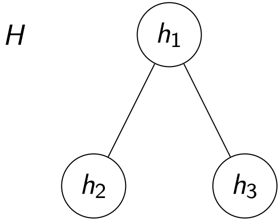


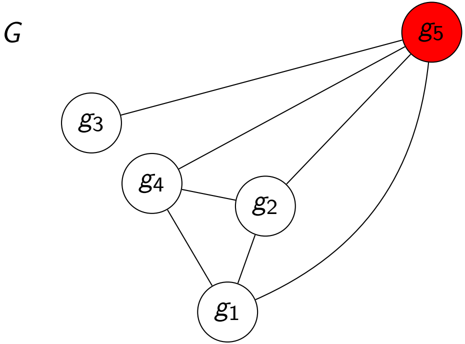


Subset of $V(H)$: $\emptyset$
Count $= 0 + 1$

$h_1 \to g_1$
$h_2 \to g_2$
$h_3 \to g_4$

Example

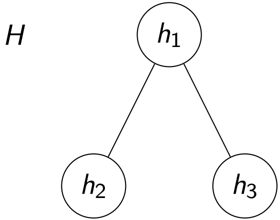


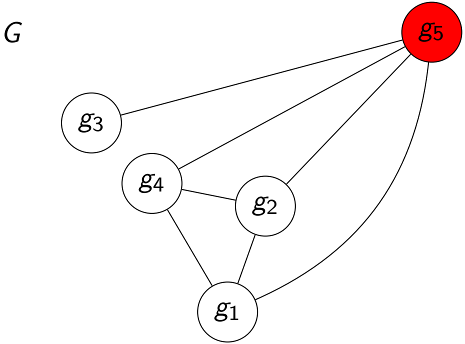


Subset of $V(H)$: $\emptyset$
Count $= 1$

$h_1 \rightarrow g_1$
$h_2 \rightarrow g_4$
$h_3 : g_2, \cancel{g_4}$

Example



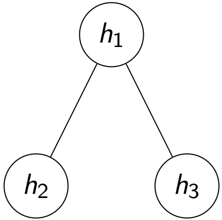


Subset of $V(H)$: $\emptyset$
Count $= 1+1$

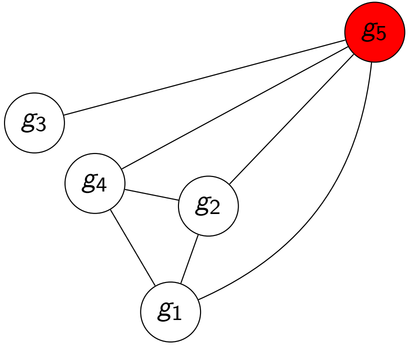
$h_1 \to g_1$
$h_2 \to g_4$
$h_3 \to g_2$

Example

H



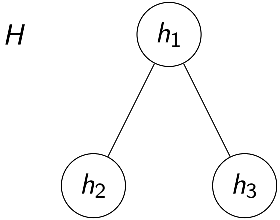
G

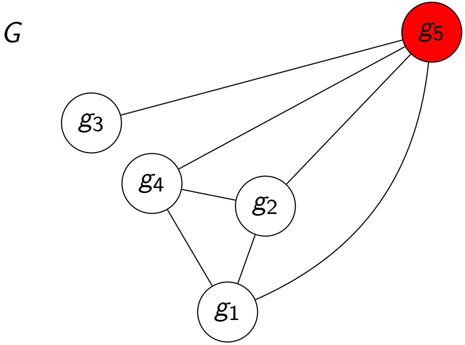


Subset of $V(H)$: $\emptyset$
Count $= 2$

$h_1 \to g_2$
$h_2 : g_1, \cancel{g_2}, \cancel{g_3}, g_4$
$h_3 : g_1, \cancel{g_2}, \cancel{g_3}, g_4$

Example

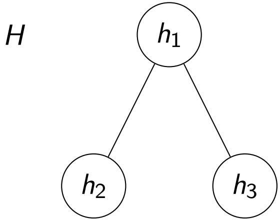


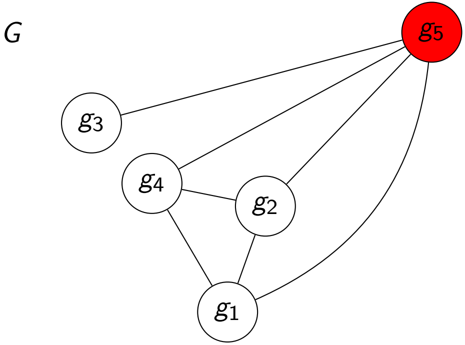


Subset of $V(H)$: $\emptyset$
Count $= 2$

$h_1 \rightarrow g_2$
$h_2 \rightarrow g_1$
$h_3 : \cancel{g_1}, g_4$

Example

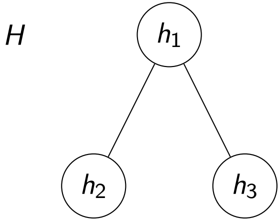


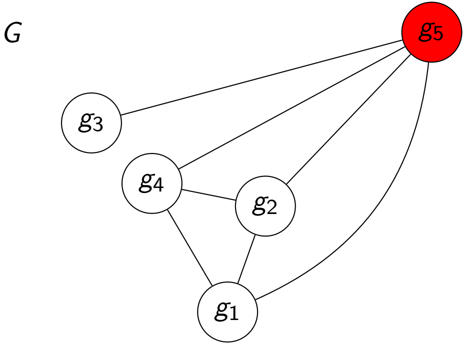


Subset of $V(H)$: $\emptyset$
Count $= 2+1$

$h_1 \to g_2$
$h_2 \to g_1$
$h_3 \to g_4$

Example

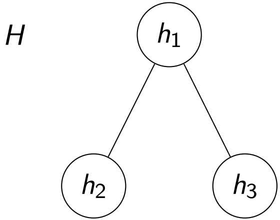


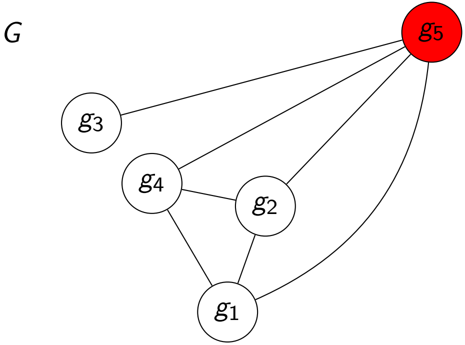


Subset of $V(H)$: $\emptyset$
Count $= 3$

$h_1 \rightarrow g_2$
$h_2 \rightarrow g_4$
$h_3 : g_1, \cancel{g_4}$

Example



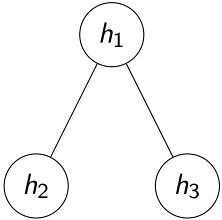


Subset of $V(H)$: $\emptyset$
Count $= 3+1$

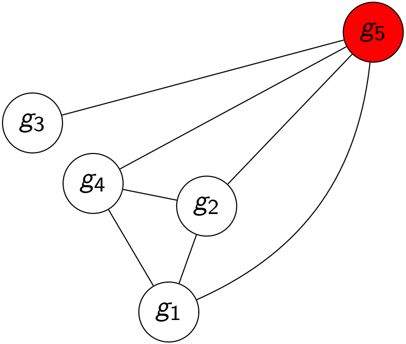
$h_1 \to g_2$
$h_2 \to g_4$
$h_3 \to g_1$

Example

H



G

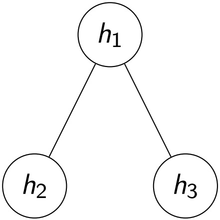


Subset of $V(H)$: $\emptyset$
Count $= 4$

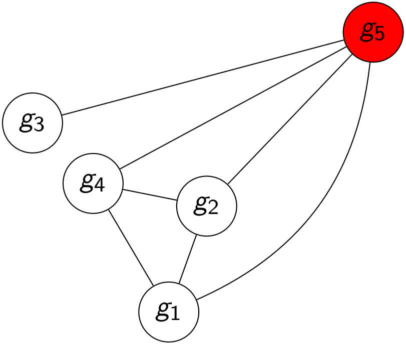
$h_1 \rightarrow g_3$
$h_2 : \cancel{g_1}, \cancel{g_2}, \cancel{g_3}, \cancel{g_4}$
$h_3 : \cancel{g_1}, \cancel{g_2}, \cancel{g_3}, \cancel{g_4}$

Example

H



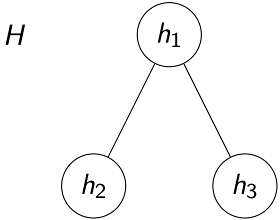
G

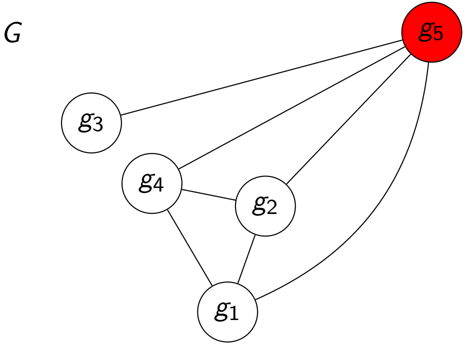


Subset of $V(H)$: $\emptyset$
Count $= 4$

$h_1 \to g_4$
h_2: $g_1, g_2, \cancel{g_3}, \cancel{g_4}$
h_3: $g_1, g_2, \cancel{g_3}, \cancel{g_4}$

Example

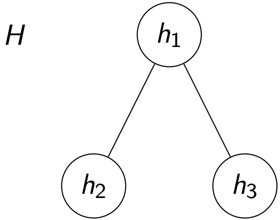


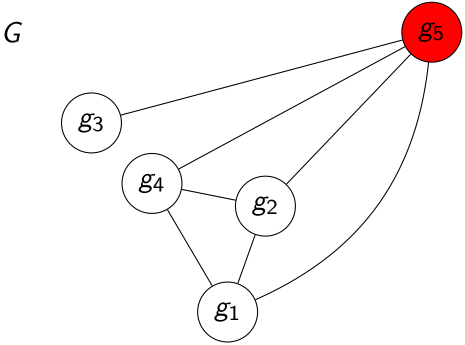


Subset of $V(H)$: $\emptyset$
Count $= 4$

$h_1 \rightarrow g_4$
$h_2 \rightarrow g_1$
h_3 : $\cancel{g_1}, g_2$

Example

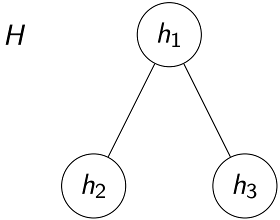


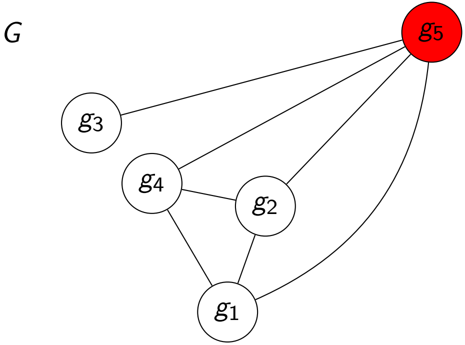


Subset of $V(H)$: $\emptyset$
Count $= 4+1$

$h_1 \to g_4$
$h_2 \to g_1$
$h_3 \to g_2$

Example

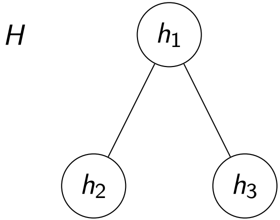


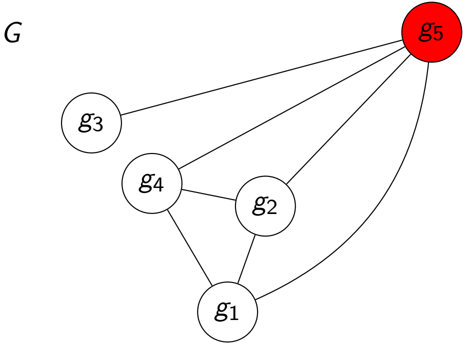


Subset of $V(H)$: $\emptyset$
Count $= 5$

$h_1 \to g_4$
$h_2 \to g_2$
$h_3 : g_1, \cancel{g_2}$

Example



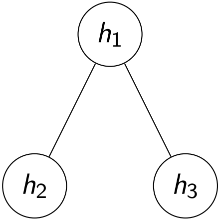


Subset of $V(H)$: $\emptyset$
Count $= 5+1$

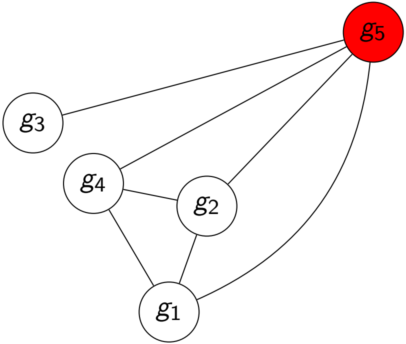
$h_1 \to g_4$
$h_2 \to g_2$
$h_3 \to g_1$

Example

H



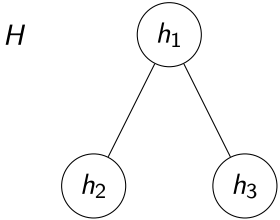
G

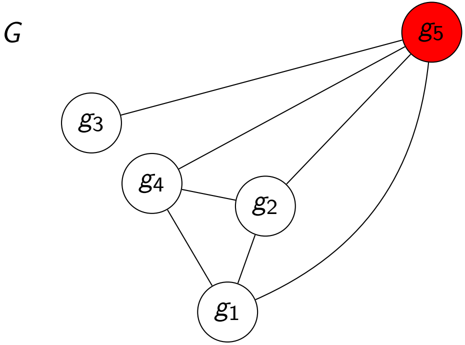


Subset of $V(H)$: h_1
Count $= 6$

$h_1 : \cancel{g_1}, \cancel{g_2}, \cancel{g_3}, \cancel{g_4}, g_5$
$h_2 : g_1, g_2, g_3, g_4, \cancel{g_5}$
$h_3 : g_1, g_2, g_3, g_4, \cancel{g_5}$

Example





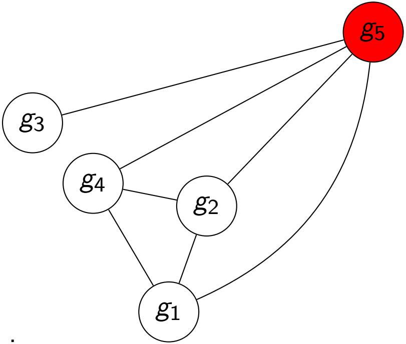
Subset of $V(H)$: h_1
Count $= 6$

$h_1 \to g_5$
$h_2 : g_1, g_2, g_3, g_4$
$h_3 : g_1, g_2, g_3, g_4$

Example

$H \setminus h_1$

G



Connected components of $H \setminus h_1$:
$C_1 = h_2$
$C_2 = h_3$

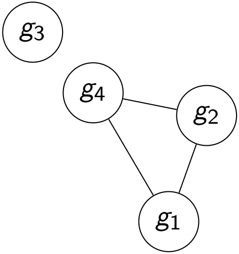
Example

$$
\begin{aligned}
\text{copies of } C_1 \text{ and } C_2 \text{ in } G \setminus g_5 = {} & (\text{copies of } C_1 \text{ in } G \setminus g_5 \\
& \times \text{copies of } C_1 \text{ in } G \setminus g_5) \\
& - \text{overlapping copies of } C_1 \text{ and } C_2 \text{ in } G \setminus g_5
\end{aligned}
$$

Counting copies of C_1 in $G \setminus g_5$

C_1

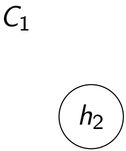
$G \setminus g_5$

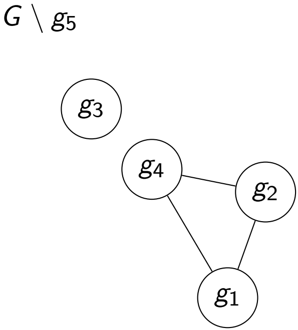


Count $=0$

$h_2: g_1, g_2, g_3, g_4, g_5$

Counting copies of C_1 in $G \setminus g_5$





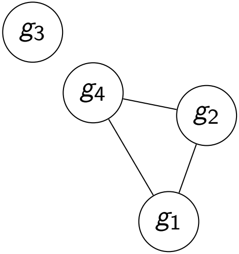
Count $= 0 + 1$

$h_2 \rightarrow g_1$

Counting copies of C_1 in $G \setminus g_5$

C_1

$G \setminus g_5$



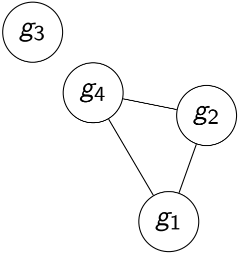
Count $= 1 + 1$

$h_2 \to g_2$

Counting copies of C_1 in $G \setminus g_5$

C_1

$G \setminus g_5$



Count $=2+1$

$h_2 \to g_3$

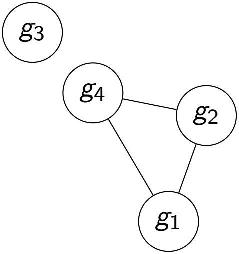
Counting copies of C_1 in $G \setminus g_5$

C_1

Count $= 3 + 1$

$h_2 \to g_4$

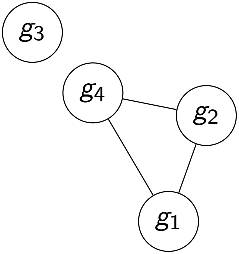
$G \setminus g_5$



Counting copies of C_1 in $G \setminus g_5$

C_1

$G \setminus g_5$



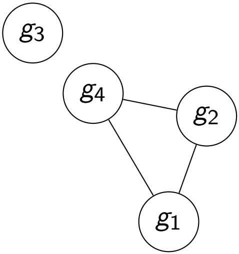
Count $=0$

$h_2: \; g_1, g_2, g_3, g_4, g_5$

Counting copies of C_2 in $G \setminus g_5$

C_2

$G \setminus g_5$



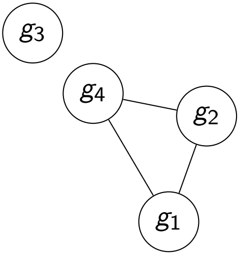
Count $=0+1$

$h_3 \to g_1$

Counting copies of C_2 in $G \setminus g_5$

C_2

$G \setminus g_5$



Count $=1+1$

$h_2 \to g_2$

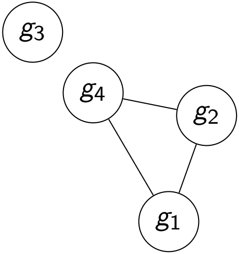
Counting copies of C_2 in $G \setminus g_5$

C_2

Count $=2$$+1$

$h_3 \to g_3$

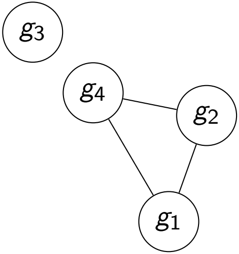
$G \setminus g_5$



Counting copies of C_2 in $G \setminus g_5$

C_2

$G \setminus g_5$

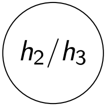


Count $= 3 + 1$

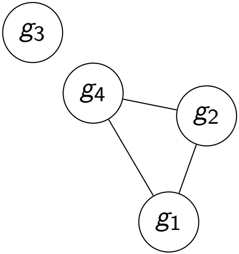
$h_3 \to g_4$

Counting overlapping copies of C_1 and C_2 in $G \setminus g_5$

$C_1 \cap C_2$



$G \setminus g_5$

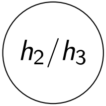


Count $=0$

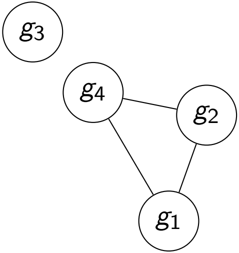
h_2/h_3 : g_1, g_2, g_3, g_4

Counting overlapping copies of C_1 and C_2 in $G \setminus g_5$

$C_1 \cap C_2$



$G \setminus g_5$

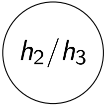


Count $= 0 + 1$

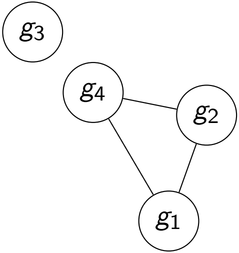
$h_2/h_3 \rightarrow g_1$

Counting overlapping copies of C_1 and C_2 in $G \setminus g_5$

$C_1 \cap C_2$



$G \setminus g_5$

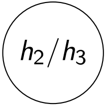


Count $=1+1$

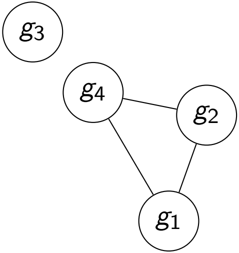
$h_2/h_3 \rightarrow g_2$

Counting overlapping copies of C_1 and C_2 in $G \setminus g_5$

$C_1 \cap C_2$



$G \setminus g_5$

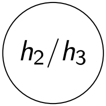


Count $= 2 + 1$

$h_2/h_3 \rightarrow g_3$

Counting overlapping copies of C_1 and C_2 in $G \setminus g_5$

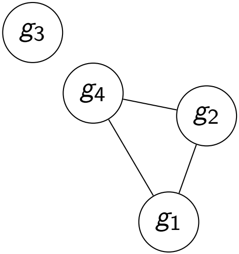
$C_1 \cap C_2$



Count $= 3 + 1$

$h_2/h_3 \to g_4$

$G \setminus g_5$



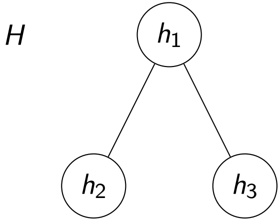
Example

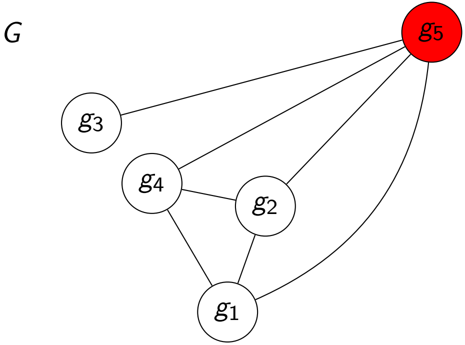
$$\begin{aligned} \text{copies of } C_1 \text{ and } C_2 \text{ in } G \setminus g_5 = &(\text{copies of } C_1 \text{ in } G \setminus g_5 \\ &\times \text{copies of } C_1 \text{ in } G \setminus g_5) \\ &- \text{overlapping copies of } C_1 \text{ and } C_2 \text{ in } G \setminus g_5 \end{aligned}$$

Example

$$\begin{aligned}\text{copies of } C_1 \text{ and } C_2 \text{ in } G \setminus g_5 &= (4 \times 4) - 4 \\ &= 12\end{aligned}$$

Example

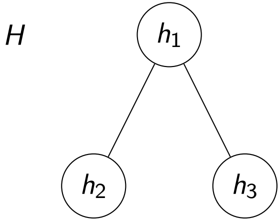


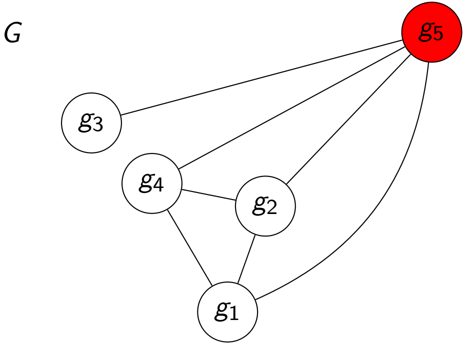


Subset of $V(H)$: h_1
Count $= 6+12$

$h_1 \to g_5$
$h_2 : g_1, g_2, g_3, g_4$
$h_3 : g_1, g_2, g_3, g_4$

Example



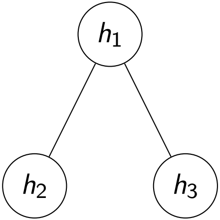


Subset of $V(H)$: h_2
Count $= 18$

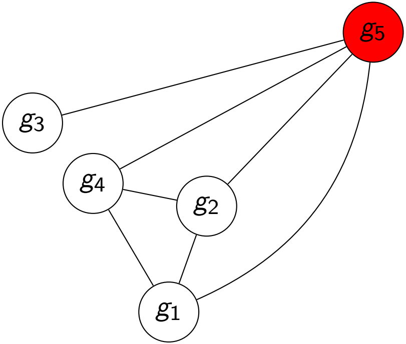
$h_1 : g_1, g_2, g_3, g_4, \cancel{g_5}$
$h_2 : \cancel{g_1}, \cancel{g_2}, \cancel{g_3}, \cancel{g_4}, g_5$
$h_3 : g_1, g_2, g_3, g_4, \cancel{g_5}$

Example

H



G

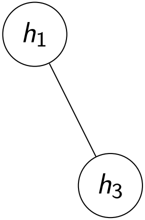


Subset of $V(H)$: h_2
Count $= 18$

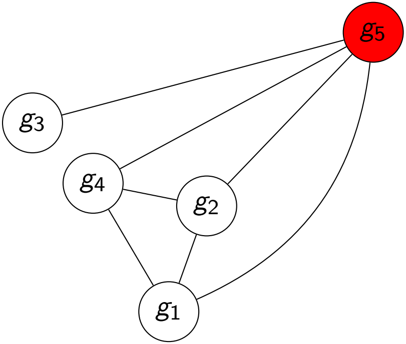
$h_1 : g_1, g_2, g_3, g_4$
$h_2 \to g_5$
$h_3 : g_1, g_2, g_3, g_4$

Example

$H \setminus h_2$



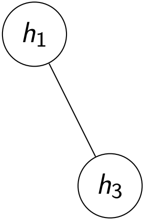
G



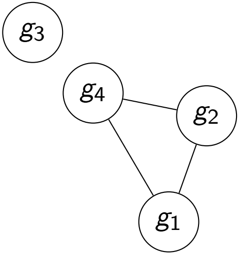
Connected components of $H \setminus h_2$:
$C_1 = h_1, h_3$

Counting copies of C_1 in $G \setminus g_5$

$H \setminus h_2$



$G \setminus g_5$

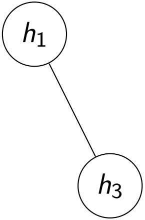


Count $=0$

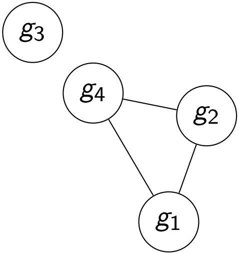
$h_1 : g_1, g_2, g_3, g_4$
$h_3 : g_1, g_2, g_3, g_4$

Counting copies of C_1 in $G \setminus g_5$

C_1



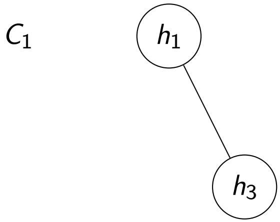
$G \setminus g_5$

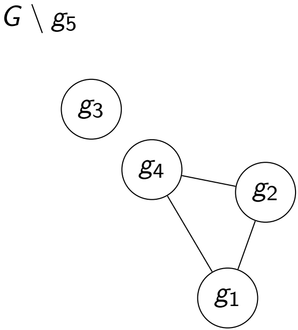


Count $=0$

$h_1 \to g_1$
$h_3 : \cancel{g_1}, g_2, \cancel{g_3}, g_4$

Counting copies of C_1 in $G \setminus g_5$

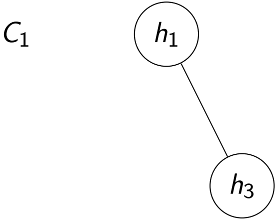


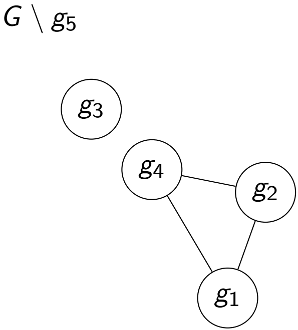


Count $= 0 + 1$

$h_1 \to g_1$
$h_3 \to g_2$

Counting copies of C_1 in $G \setminus g_5$



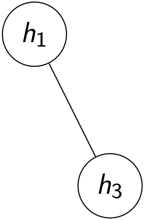


Count $=1+1$

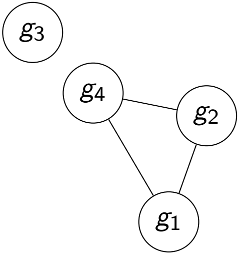
$h_1 \to g_1$
$h_3 \to g_4$

Counting copies of C_1 in $G \setminus g_5$

C_1



$G \setminus g_5$

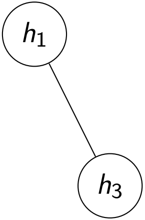


Count $=2$

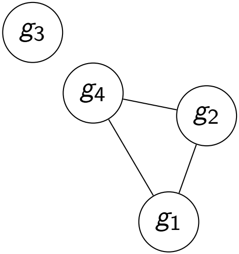
$h_1 \rightarrow g_2$
$h_3 : g_1, \cancel{g_2}, \cancel{g_3}, g_4$

Counting copies of C_1 in $G \setminus g_5$

C_1



$G \setminus g_5$

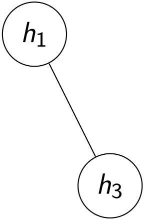


Count $=2+1$

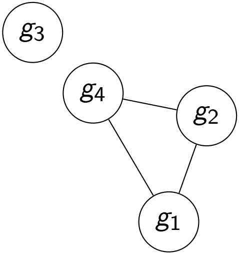
$h_1 \to g_2$
$h_3 \to g_1$

Counting copies of C_1 in $G \setminus g_5$

C_1



$G \setminus g_5$

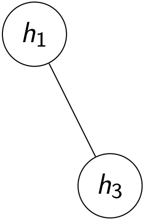


Count $= 3 + 1$

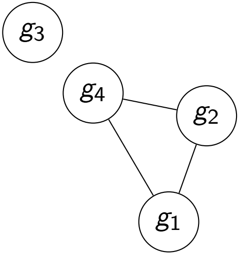
$h_1 \to g_2$
$h_3 \to g_4$

Counting copies of C_1 in $G \setminus g_5$

C_1



$G \setminus g_5$

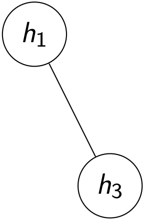


Count $=4$

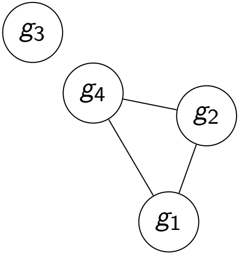
$h_1 \rightarrow g_3$
$h_3 : \cancel{g_1}, \cancel{g_2}, \cancel{g_3}, \cancel{g_4}$

Counting copies of C_1 in $G \setminus g_5$

C_1



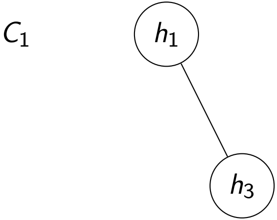
$G \setminus g_5$

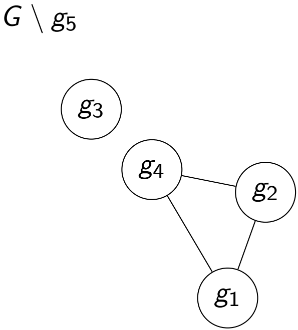


Count $=4$

$h_1 \rightarrow g_4$
$h_3 : g_1, g_2, \cancel{g_3}, \cancel{g_4}$

Counting copies of C_1 in $G \setminus g_5$



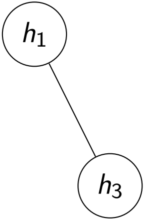


Count $=4+1$

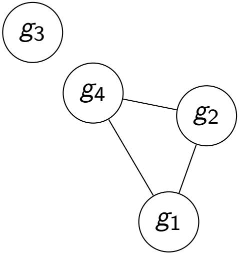
$h_1 \to g_4$
$h_3 \to g_1$

Counting copies of C_1 in $G \setminus g_5$

C_1



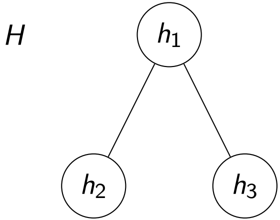
$G \setminus g_5$

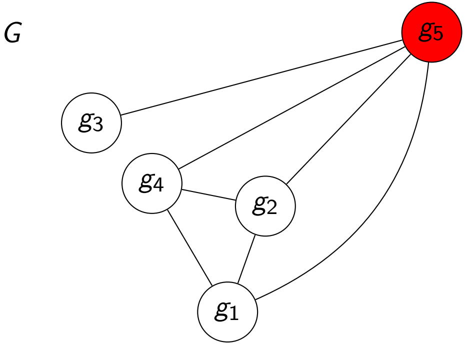


Count $=5+1$

$h_1 \rightarrow g_4$
$h_3 \rightarrow g_2$

Example

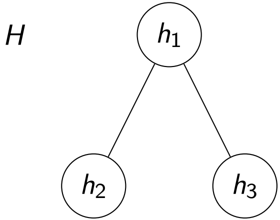


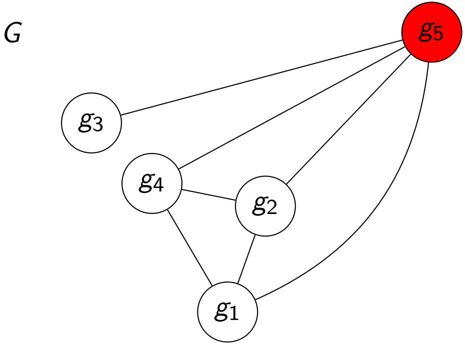


Subset of $V(H)$: h_2
Count $= 18+6$

$h_1: g_1, g_2, g_3, g_4, g_5$
$h_2: g_5$
$h_3: g_1, g_2, g_3, g_4$

Example

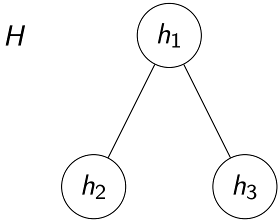


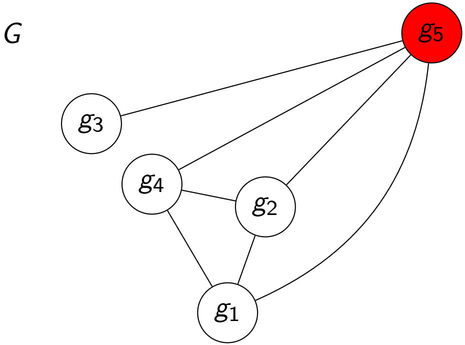


Subset of $V(H)$: h_3
Count $= 24+6$

$h_1: g_1, g_2, g_3, g_4$
$h_2: g_1, g_2, g_3, g_4$
$h_3: g_5$

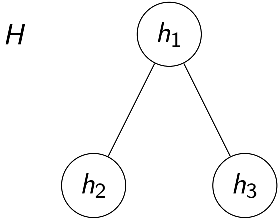
Example





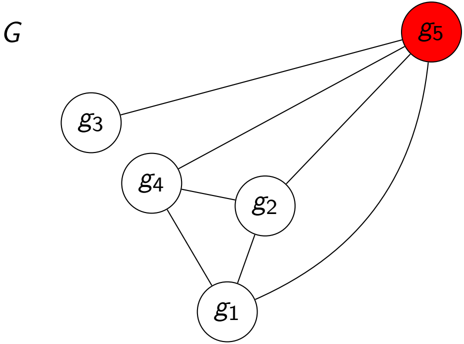
Count $= 30$

Example

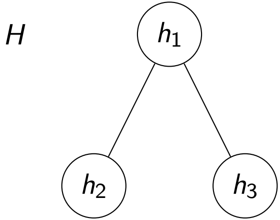


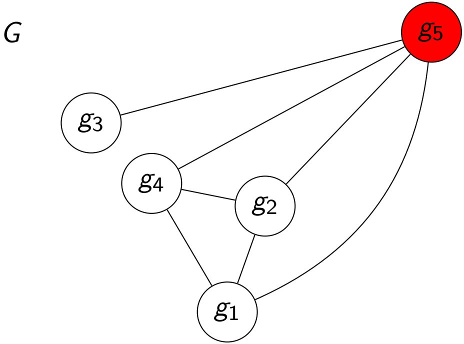
Count $= 30$

Number of copies of H in $H = 2$



Example





Count $= 30$

Number of copies of H in $H = 2$

$\rightarrow$ Number of unlabelled copies of H in $G = 30/2 = 15$