An FPT algorithm for counting subgraphs

Jess Ryan

December 30, 2018

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.

Why do we care?

- Generalisation of subgraph isomorphism
- Network analysis

Algorithmic complexity

NP-hard

NP-hard problems cannot be solved in polynomial time unless $P = NP$, which we do not expect is true.

Fixed-parameter tractable (FPT)

A problem is FPT if it can be solved in time depending polynomially on the input size and exponentially on some other parameter.

Subgraph counting is hard!

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

Bounded degree graph

A graph G has *bounded degree* k if each vertex in G has degree at most k.

Almost bounded degree graph

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

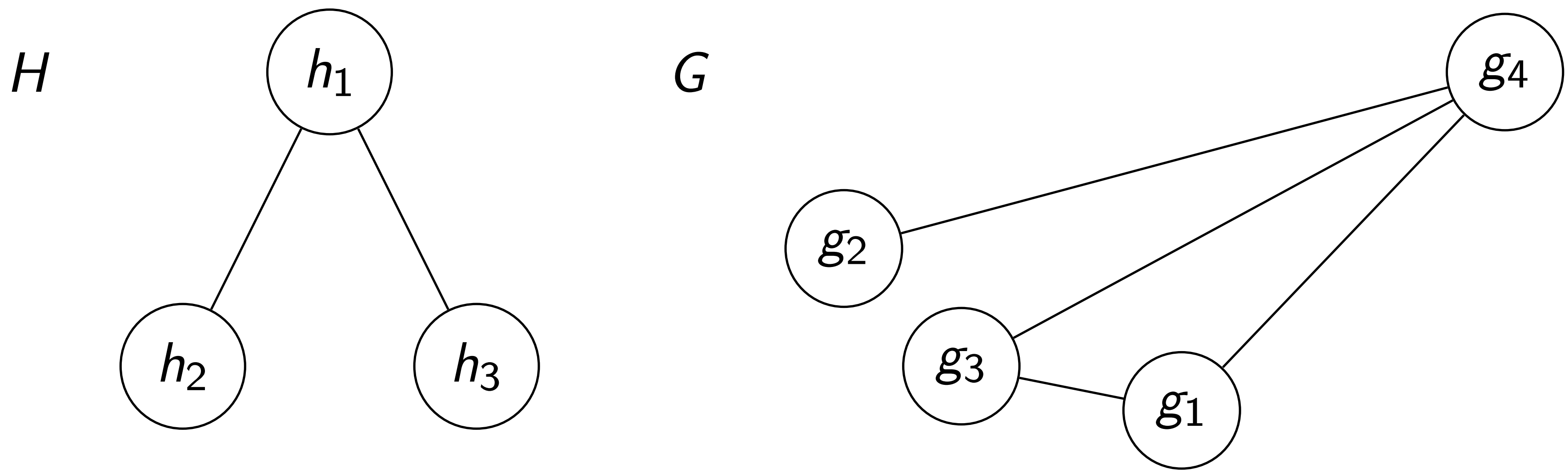
Project objective:

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

Algorithm: General Idea

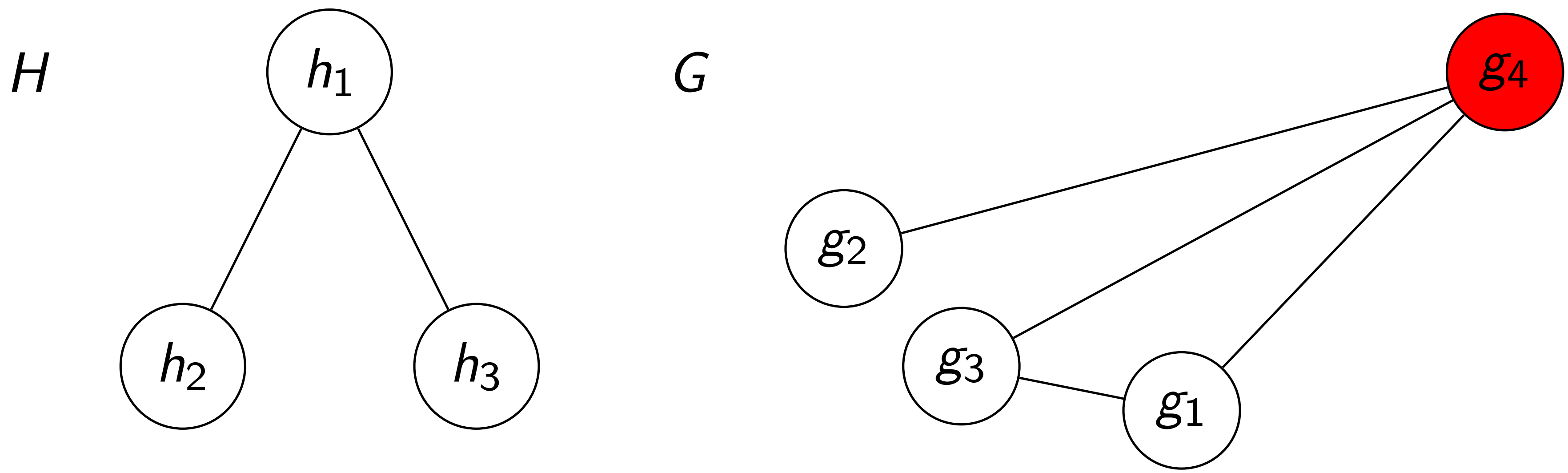
- Consider each way to assign part of H to high degree vertices in G
- For each feasible assignment, count ways to assign the rest of H to the bounded degree part of G
- Sum up the counts to obtain number of labelled copies of H in G
- Divide by the number of copies of H in H to obtain number of *unlabelled*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

Example



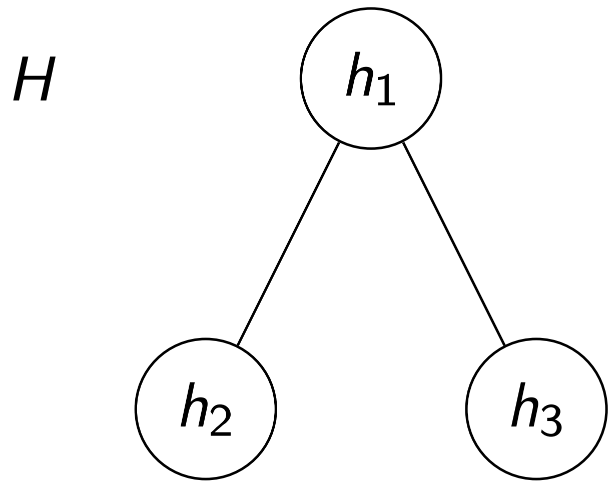
$h_1 : g_1, g_2, g_3, g_4$
$h_2 : g_1, g_2, g_3, g_4$
$h_3 : g_1, g_2, g_3, g_4$

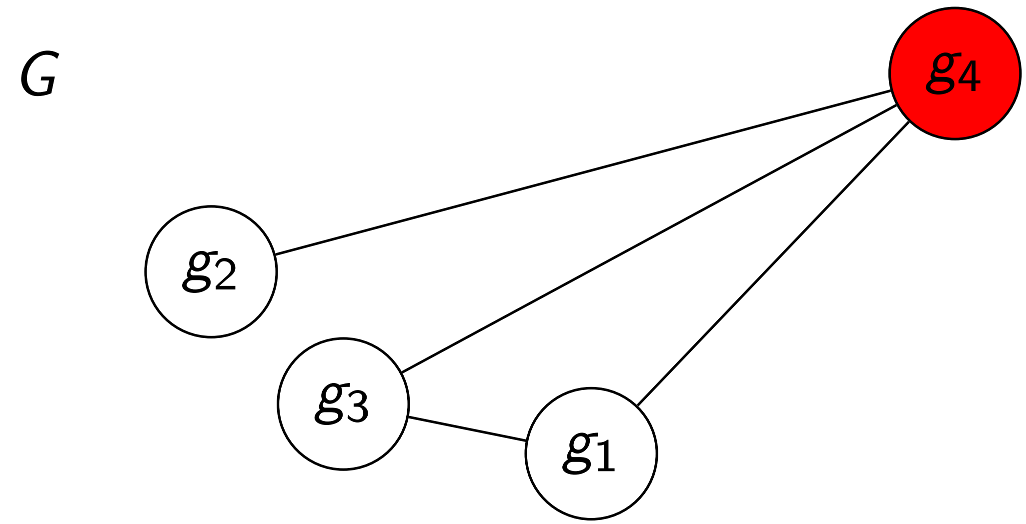
Example



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

Example

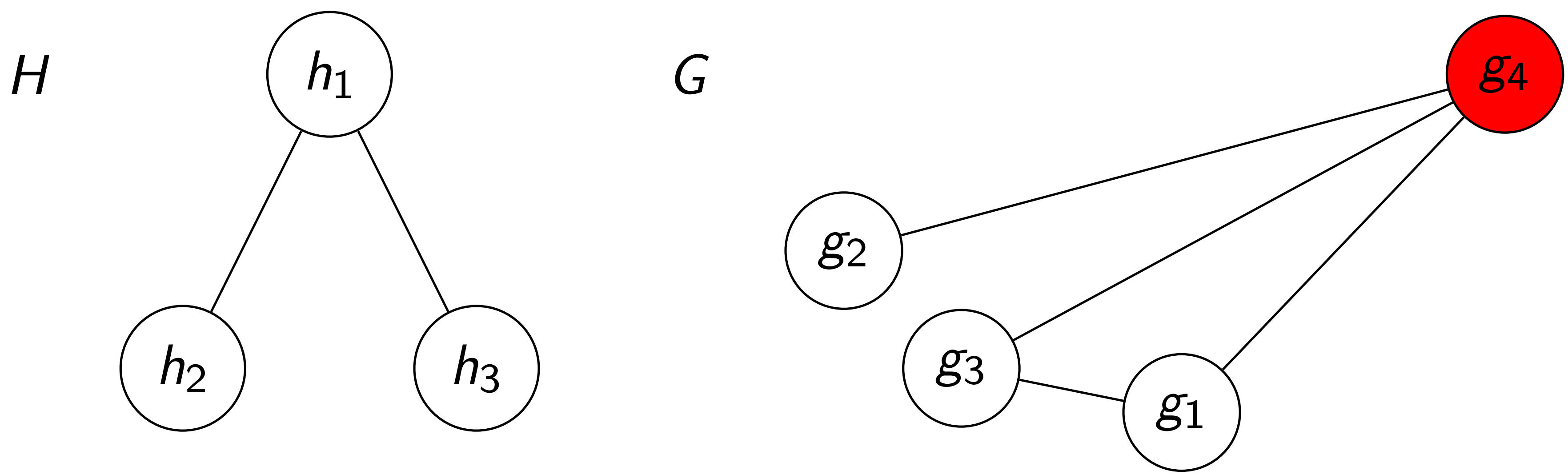




Subset of $V(H)$: $\emptyset$
Count $= 0$

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

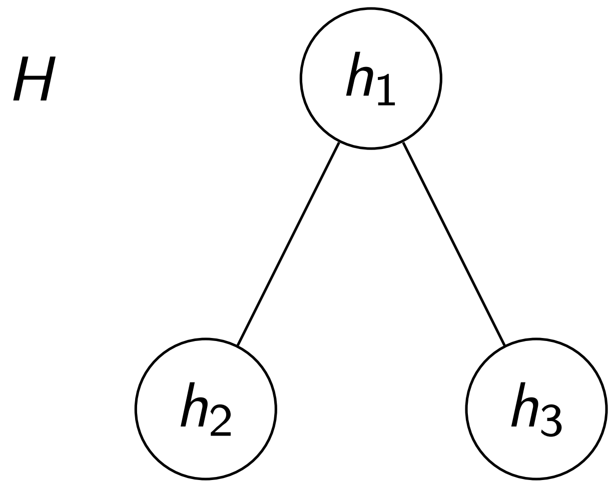
Example

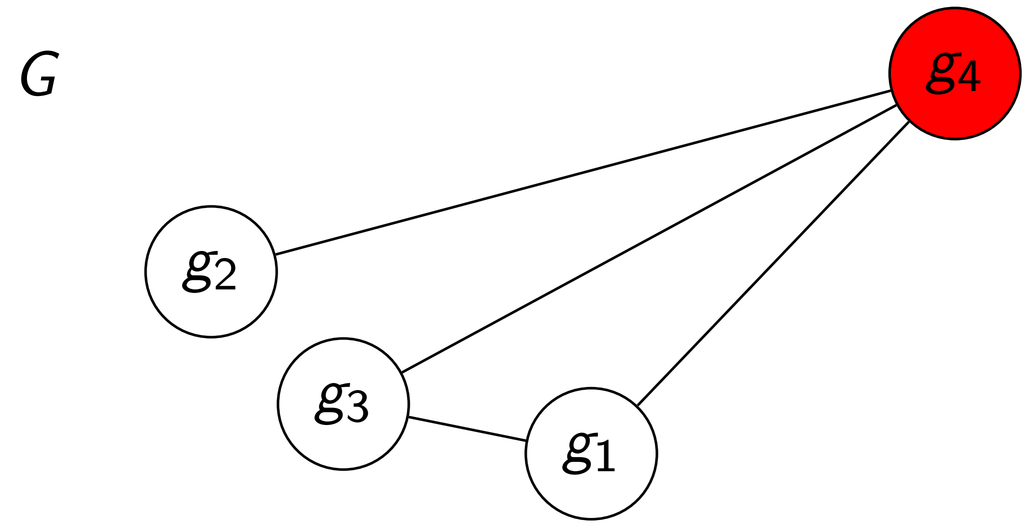


Subset of $V(H)$: $\emptyset$
Count $= 0$

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

Example





Subset of $V(H)$: $\emptyset$

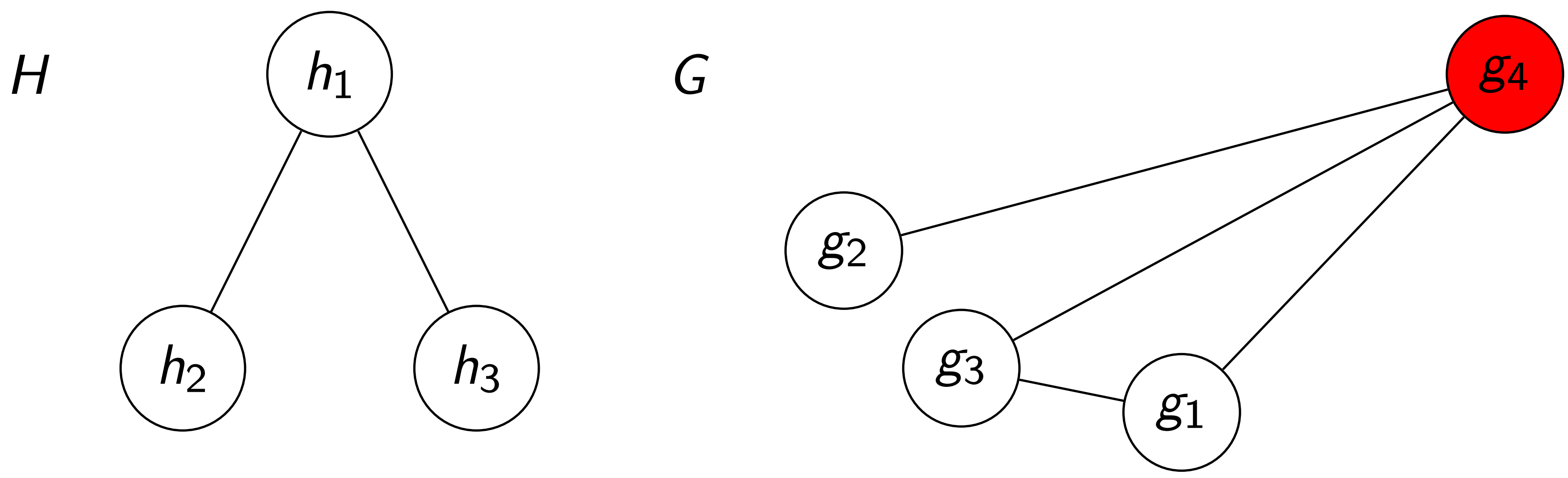
Count $= 0$

$h_1 \to g_1$

$h_2 \to g_3$

$h_3 : \cancel{g_3}$

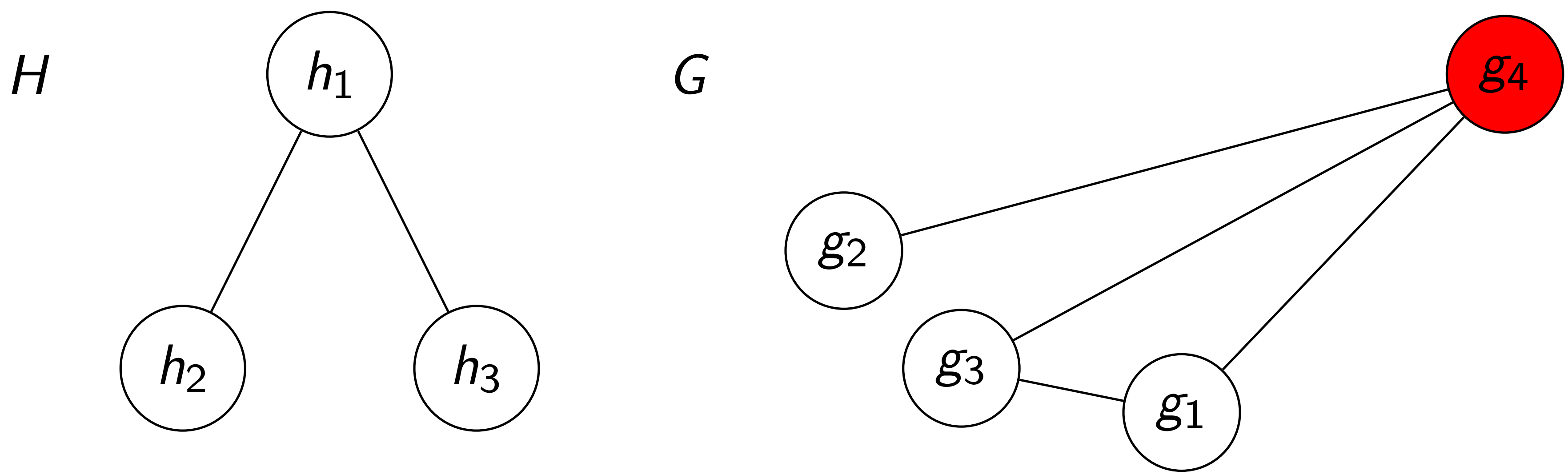
Example



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

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

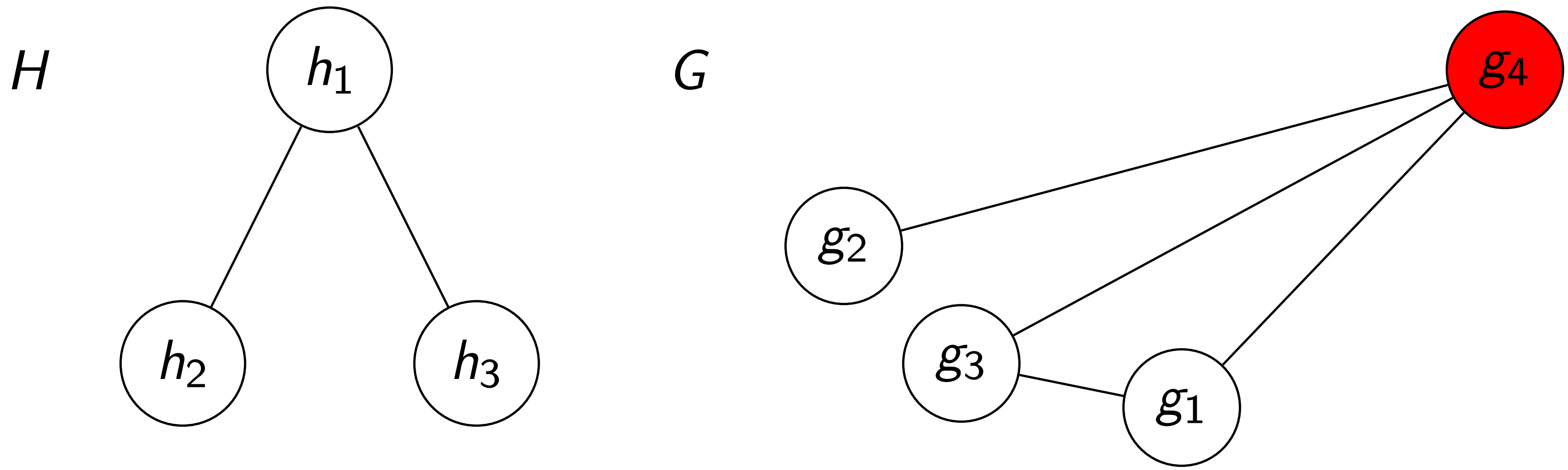
Example



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

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

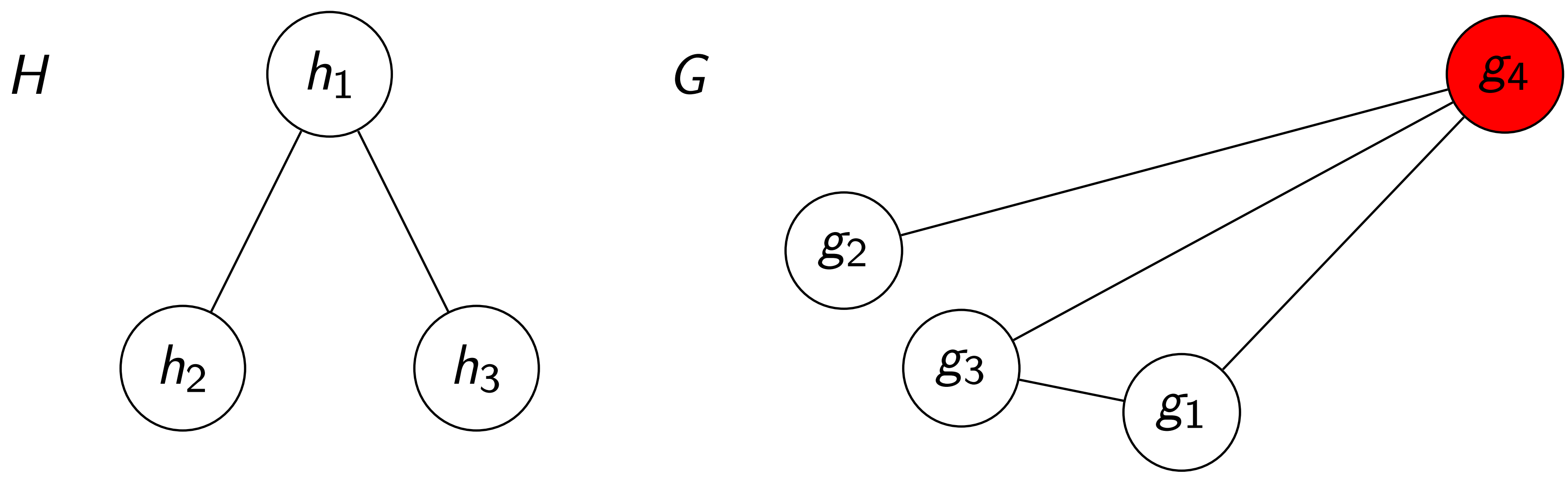
Example



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

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

Example



Subset of $V(H)$: h_1

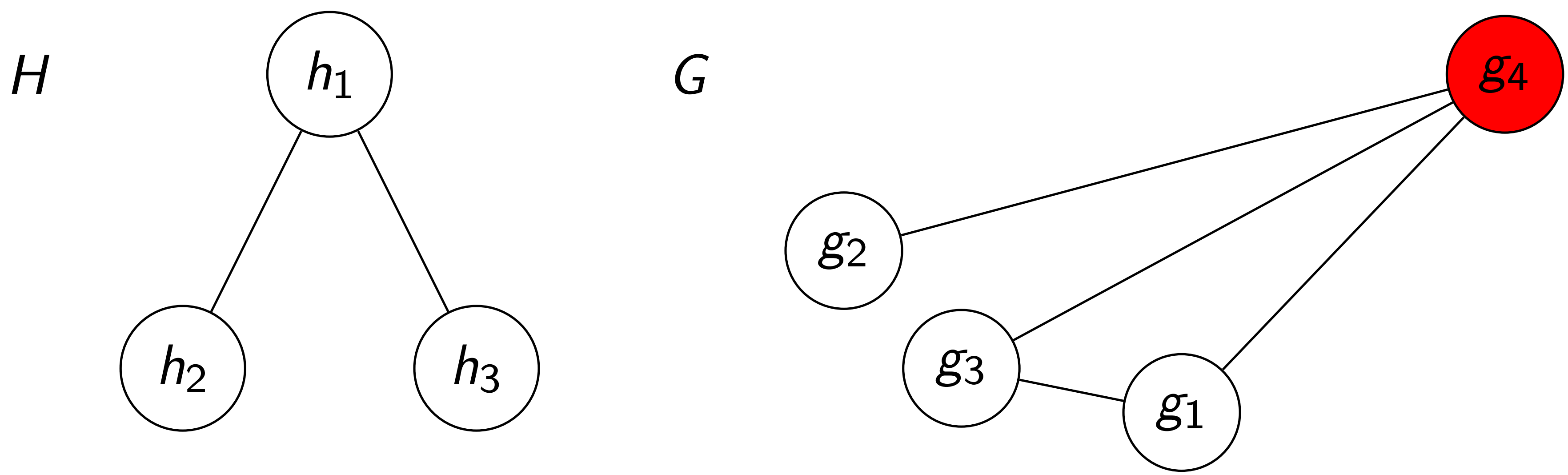
Count $= 0$

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

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

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

Example



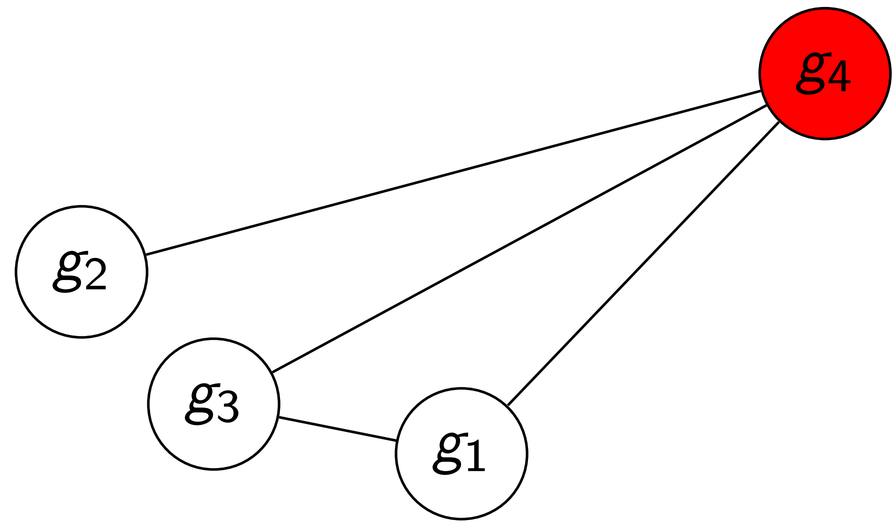
Subset of $V(H)$: h_1
Count $= 0$

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

Example

$H \setminus h_1$

G



Connected components of $H \setminus h_1$:

$C_1 = h_2$

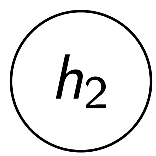
$C_2 = h_3$

Example

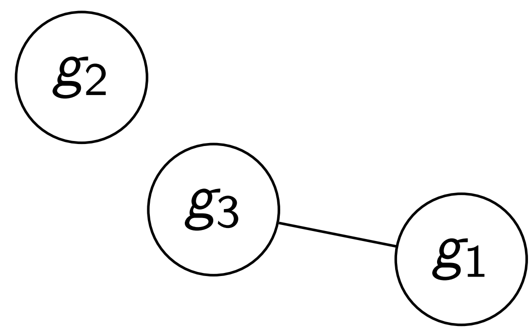
$$\begin{aligned}\text{copies of } C_1 \text{ and } C_2 \text{ in } G \setminus g_4 &= (\text{copies of } C_1 \text{ in } G \setminus g_4 \\ &\times \text{copies of } C_2 \text{ in } G \setminus g_4) \\ &- \text{overlapping copies of } C_1 \text{ and } C_2 \text{ in } G \setminus g_4\end{aligned}$$

Counting copies of C_1 in $G \setminus g_4$

C_1



$G \setminus g_4$

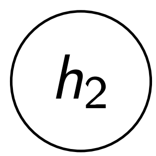


Count $=0$

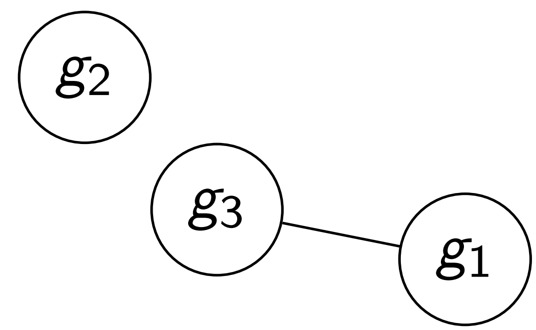
$h_2 : g_1, g_2, g_3$

Counting copies of C_1 in $G \setminus g_4$

C_1



$G \setminus g_4$

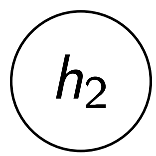


Count $=0+1$

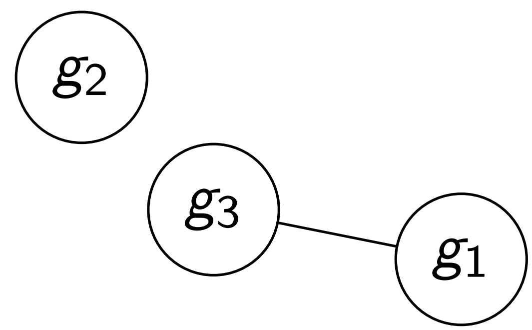
$h_2 \rightarrow g_1$

Counting copies of C_1 in $G \setminus g_4$

C_1



$G \setminus g_4$

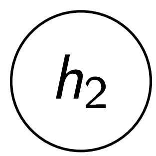


Count $= 1 + 1$

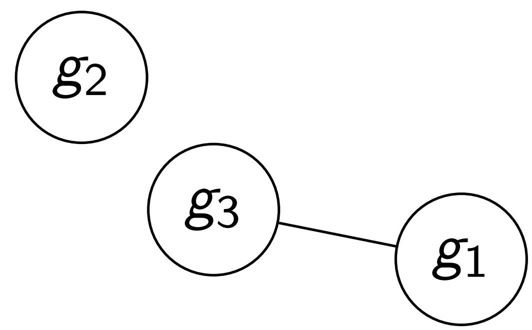
$h_2 \to g_2$

Counting copies of C_1 in $G \setminus g_4$

C_1



$G \setminus g_4$

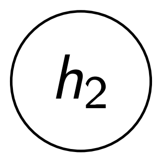


Count $=2+1=3$

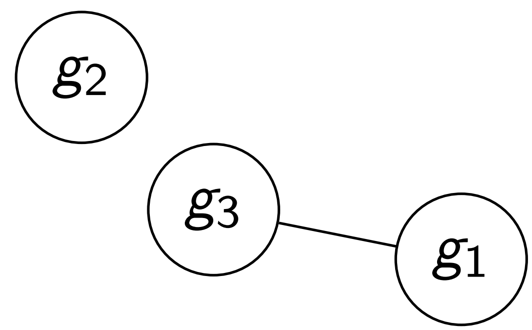
$h_2 \rightarrow g_3$

Counting copies of C_2 in $G \setminus g_4$

C_2



$G \setminus g_4$

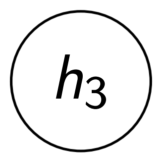


Count $=0$

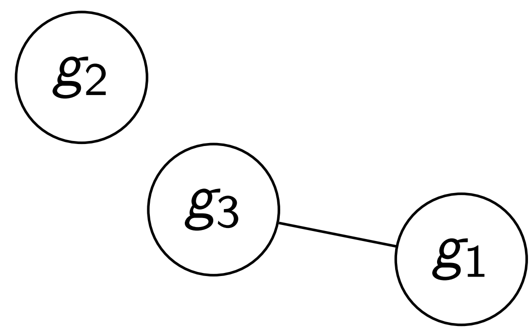
$h_3 : g_1, g_2, g_3$

Counting copies of C_2 in $G \setminus g_4$

C_2



$G \setminus g_4$

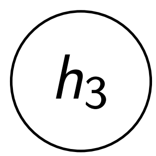


Count $=0+1$

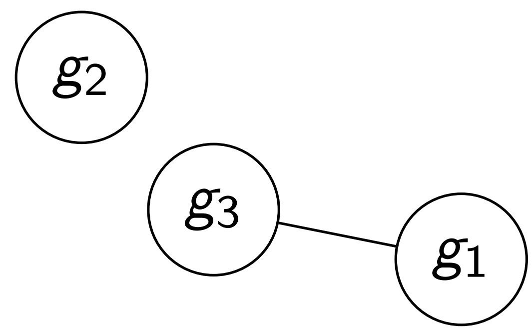
$h_3 \rightarrow g_1$

Counting copies of C_2 in $G \setminus g_4$

C_2



$G \setminus g_4$

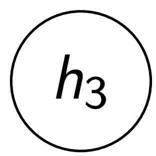


Count $= 1 + 1$

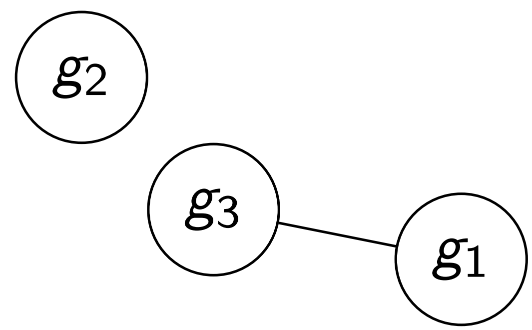
$h_3 \to g_2$

Counting copies of C_2 in $G \setminus g_4$

C_2



$G \setminus g_4$

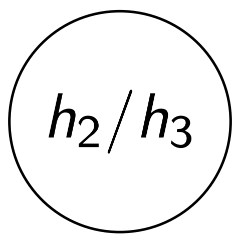


Count $=2+1=3$

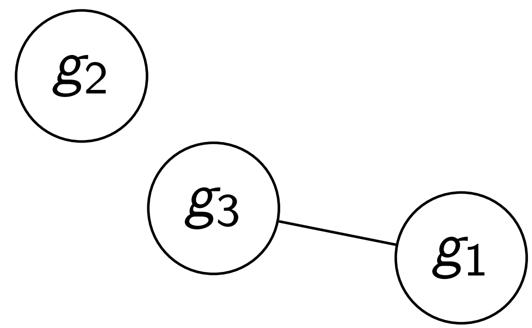
$h_3 \rightarrow g_3$

Counting overlapping copies of C_1 and C_2 in $G \setminus g_4$

$C_1 \cap C_2$



$G \setminus g_4$

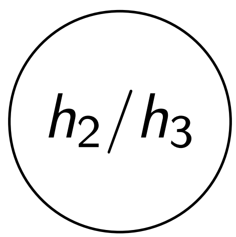


Count $=0$

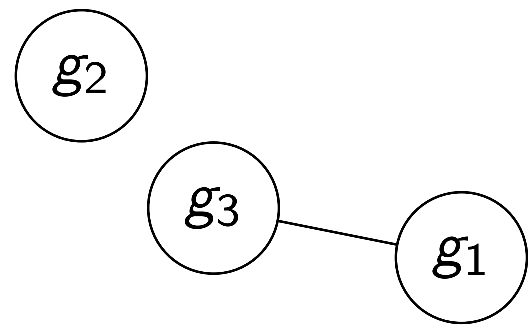
$h_2/h_3 : g_1, g_2, g_3$

Counting overlapping copies of C_1 and C_2 in $G \setminus g_4$

$C_1 \cap C_2$



$G \setminus g_4$

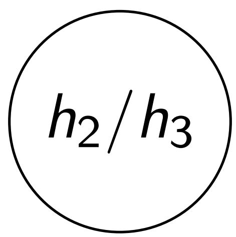


Count $=0+1$

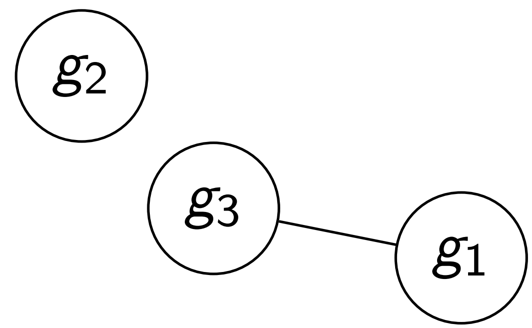
$h_2/h_3 \rightarrow g_1$

Counting overlapping copies of C_1 and C_2 in $G \setminus g_4$

$C_1 \cap C_2$



$G \setminus g_4$

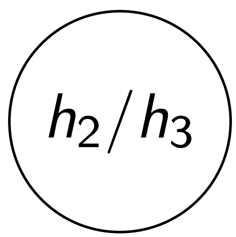


Count $= 1 + 1$

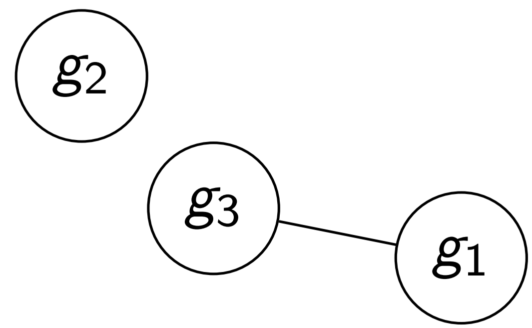
$h_2/h_3 \to g_2$

Counting overlapping copies of C_1 and C_2 in $G \setminus g_4$

$C_1 \cap C_2$



$G \setminus g_4$



Count $=2+1=3$

$h_2/h_3 \to g_3$

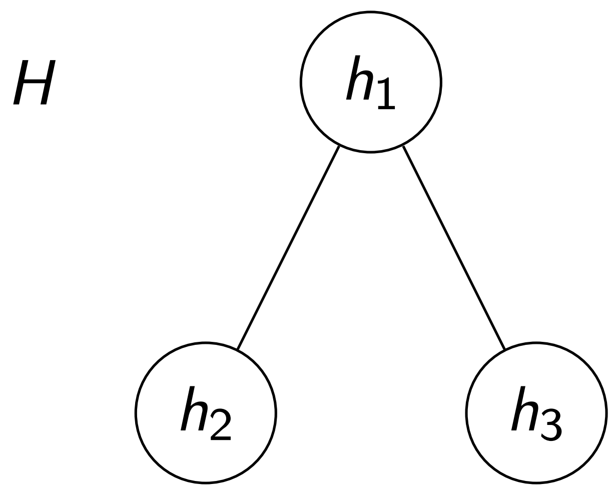
Example

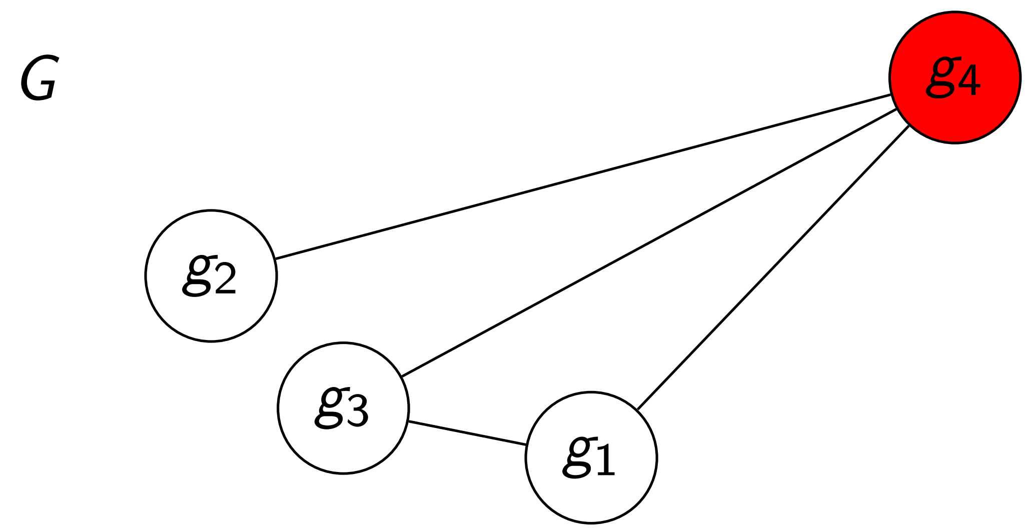
$$\begin{aligned}\text{copies of } C_1 \text{ and } C_2 \text{ in } G \setminus g_4 = (&\text{copies of } C_1 \text{ in } G \setminus g_4 \\ &\times \text{copies of } C_1 \text{ in } G \setminus g_4) \\ &- \text{overlapping copies of } C_1 \text{ and } C_2 \text{ in } G \setminus g_4\end{aligned}$$

Example

$$
\begin{aligned}
\text{copies of } C_1 \text{ and } C_2 \text{ in } G \setminus g_4 &= (3 \times 3) - 3 \\
&= 6
\end{aligned}
$$

Example

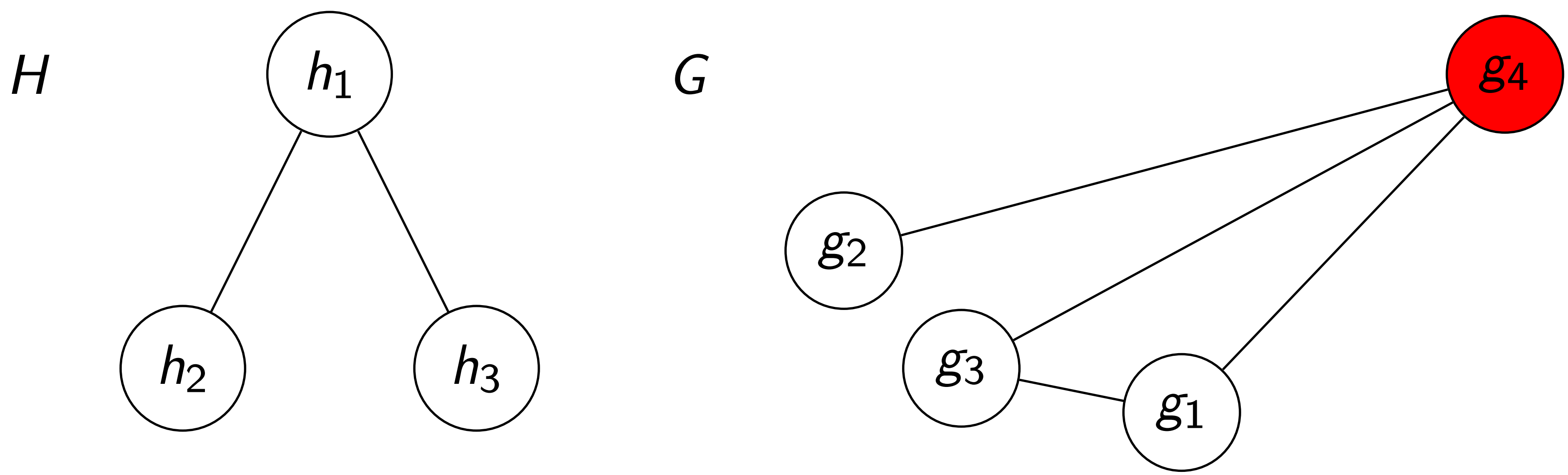




Subset of $V(H)$: h_1

Count $= 0+6$

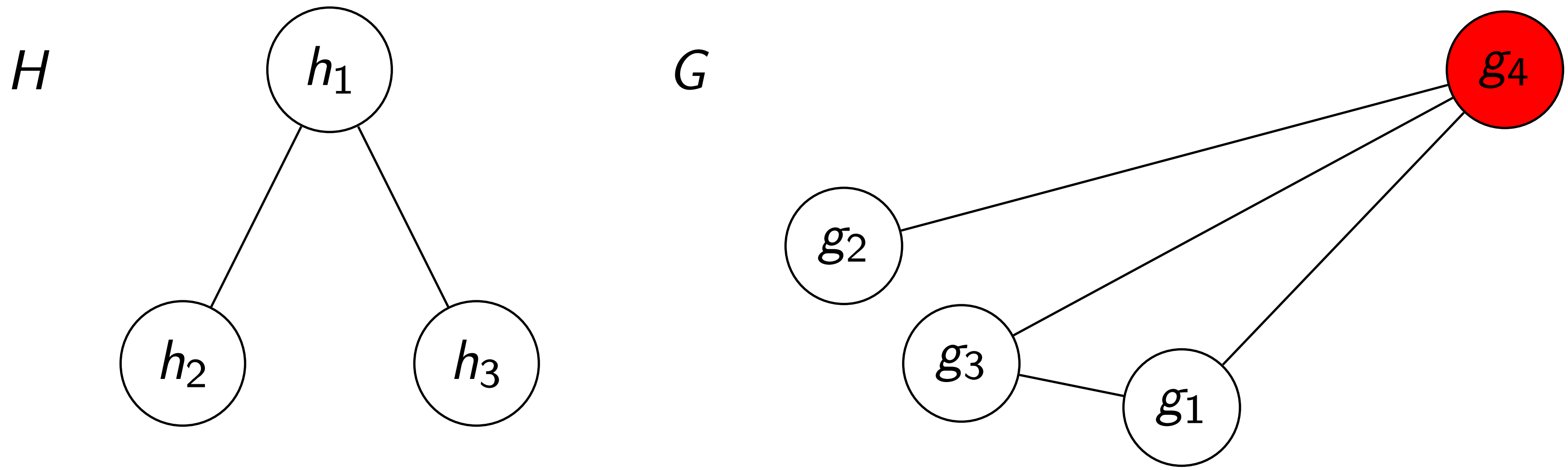
Example



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

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

Example



Subset of $V(H)$: h_2

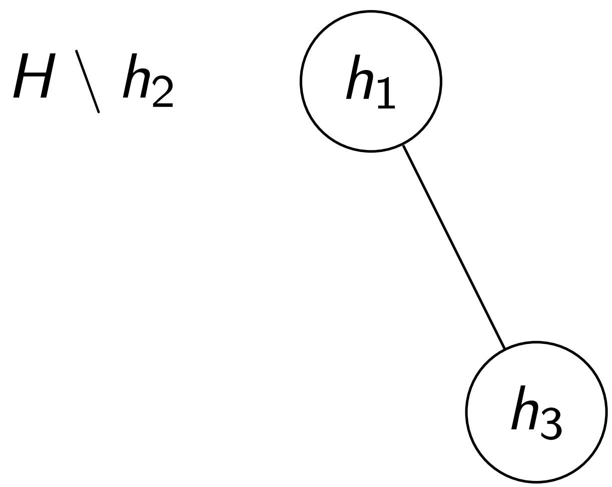
Count $= 6$

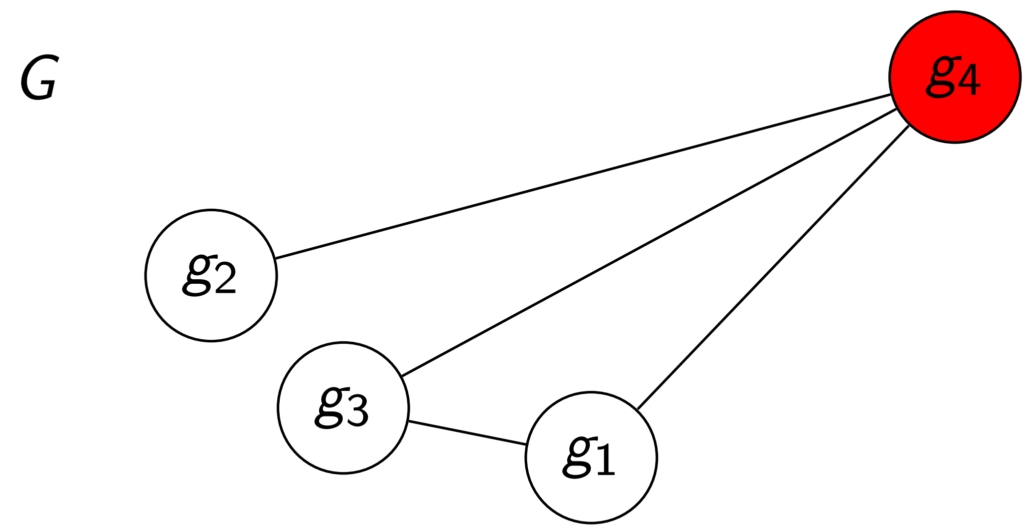
$h_1 : g_1, g_2, g_3$

$h_2 \rightarrow g_4$

$h_3 : g_1, g_2, g_3$

Example

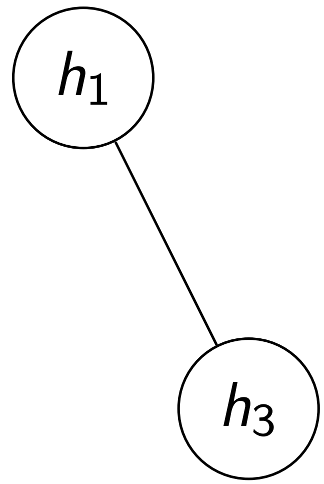




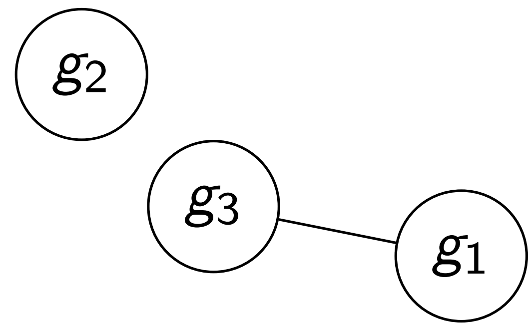
Connected components of $H \setminus h_2$:
$C_1 = h_1, h_3$

Counting copies of C_1 in $G \setminus g_4$

C_1



$G \setminus g_4$

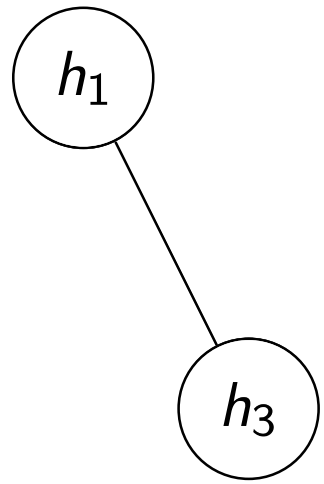


Count $=0$

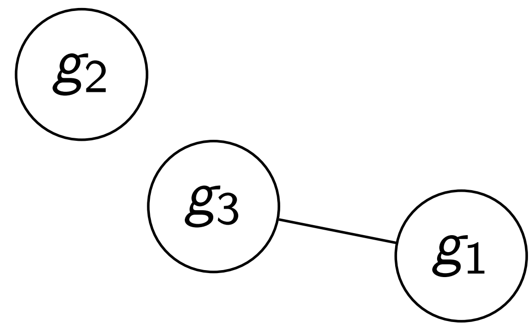
$h_1 : g_1, g_2, g_3$
$h_3 : g_1, g_2, g_3$

Counting copies of C_1 in $G \setminus g_4$

C_1



$G \setminus g_4$

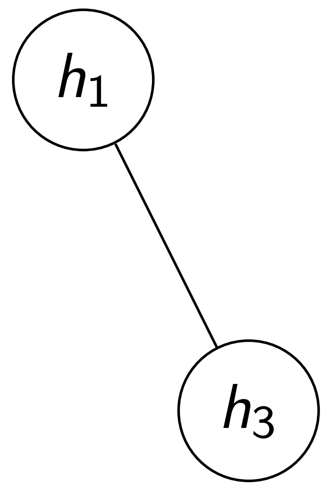


Count $=0$

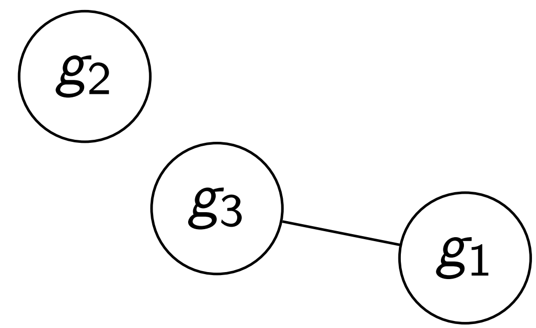
$h_1 \to g_1$
$h_3 : \cancel{g_1}, \cancel{g_2}, g_3$

Counting copies of C_1 in $G \setminus g_4$

C_1



$G \setminus g_4$

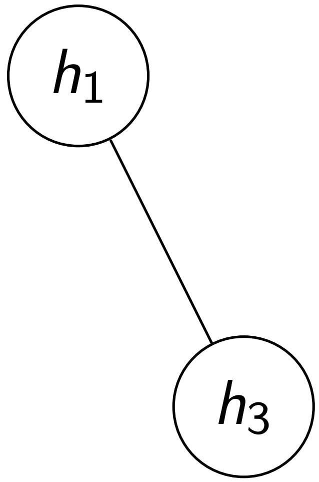


Count $=0+1$

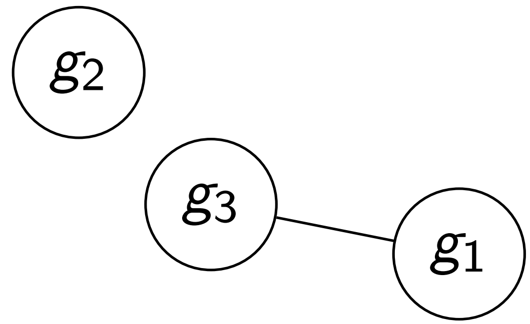
$h_1 \to g_1$
$h_3 \to g_3$

Counting copies of C_1 in $G \setminus g_4$

C_1



$G \setminus g_4$

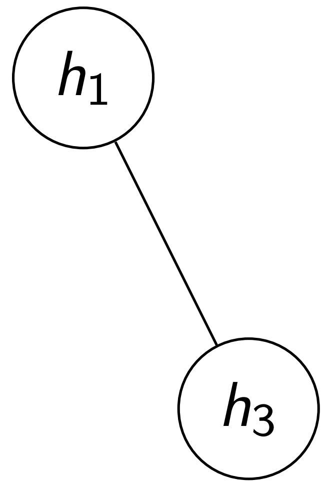


Count $=1$

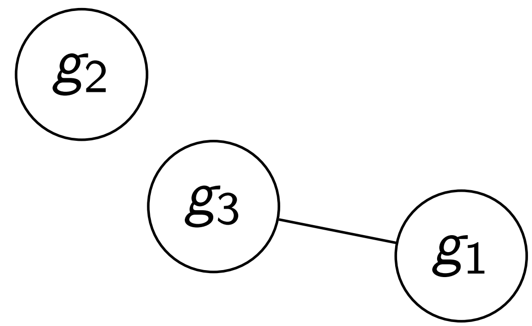
$h_1 \to g_2$
$h_3: \cancel{g_1}, \cancel{g_2}, \cancel{g_3}$

Counting copies of C_1 in $G \setminus g_4$

C_1



$G \setminus g_4$

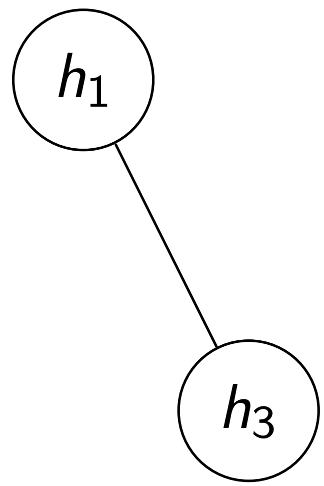


Count $=1$

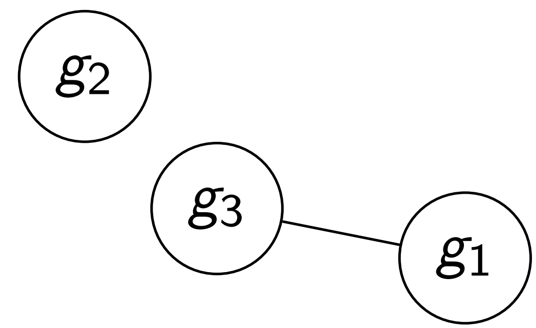
$h_1 \to g_3$
$h_3 : g_1, \cancel{g_2}, \cancel{g_3}$

Counting copies of C_1 in $G \setminus g_4$

C_1



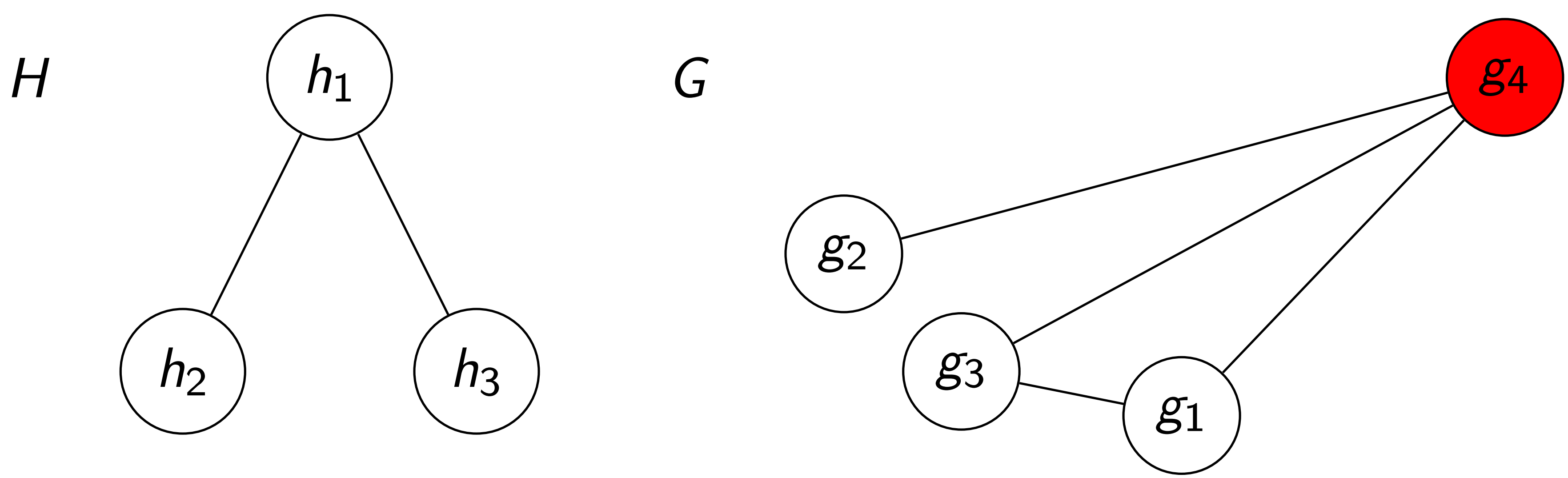
$G \setminus g_4$



Count $=1+1$

$h_1 \to g_3$
$h_3 \to g_1$

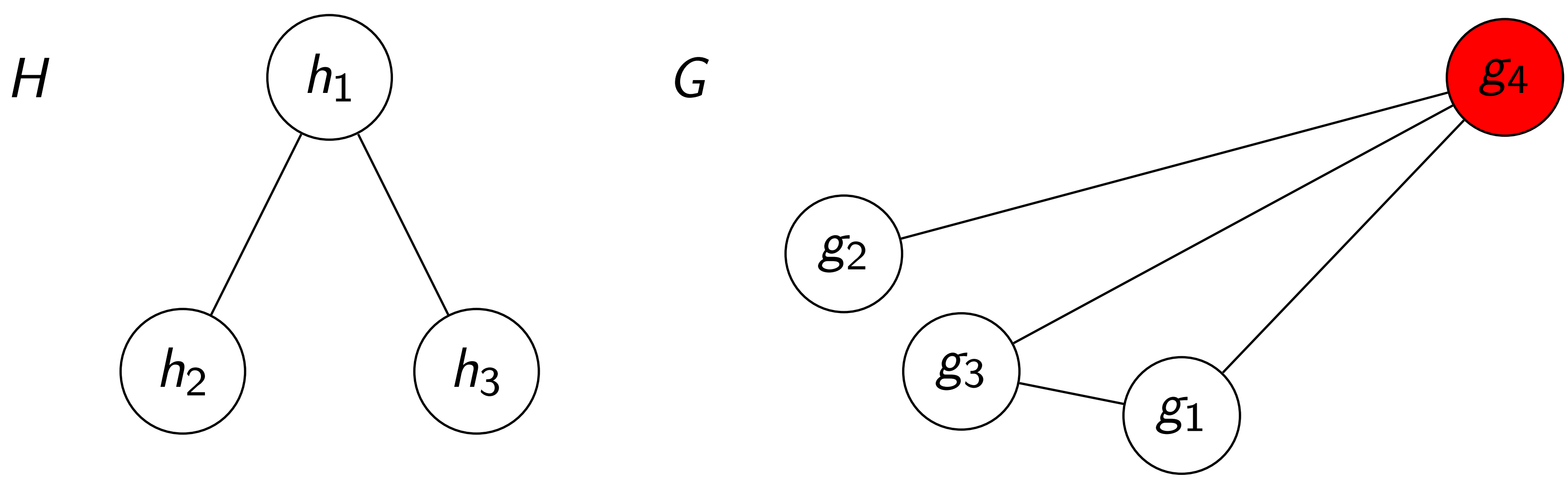
Example



Subset of $V(H)$: h_2

Count $= 6+2$

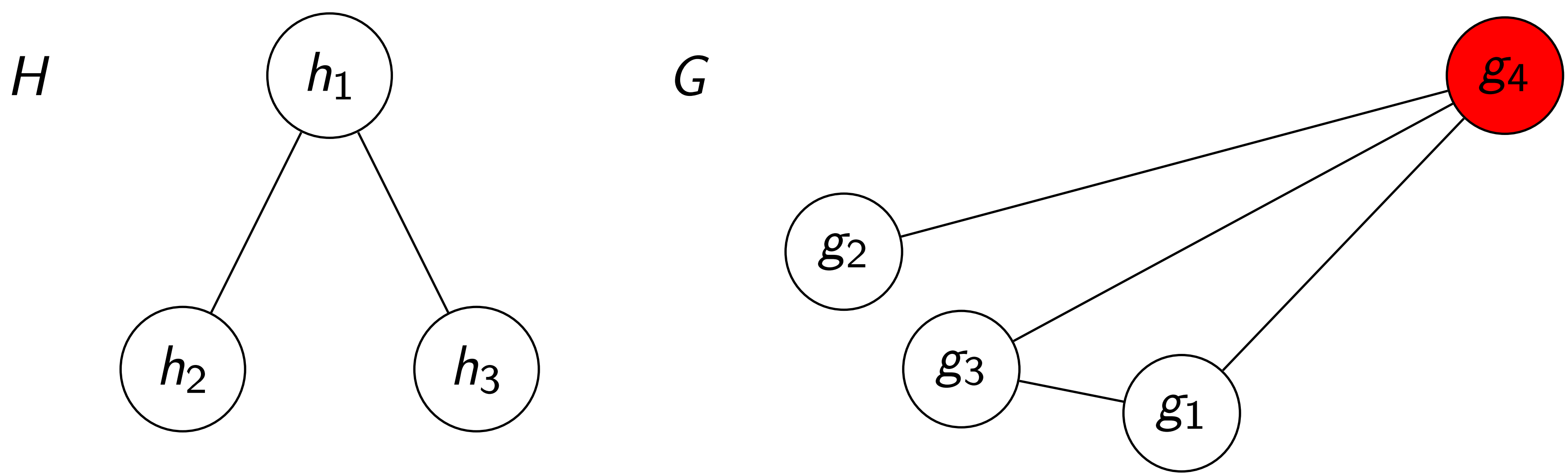
Example



Subset of $V(H)$: h_3

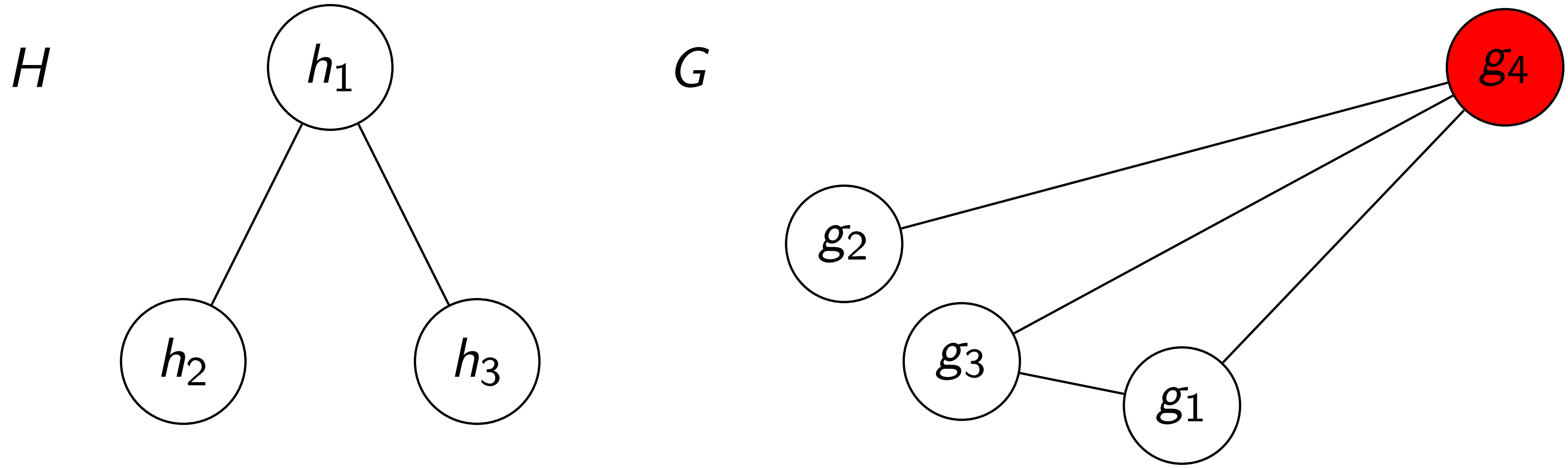
Count $= 8+2$

Example



Count $= 10$

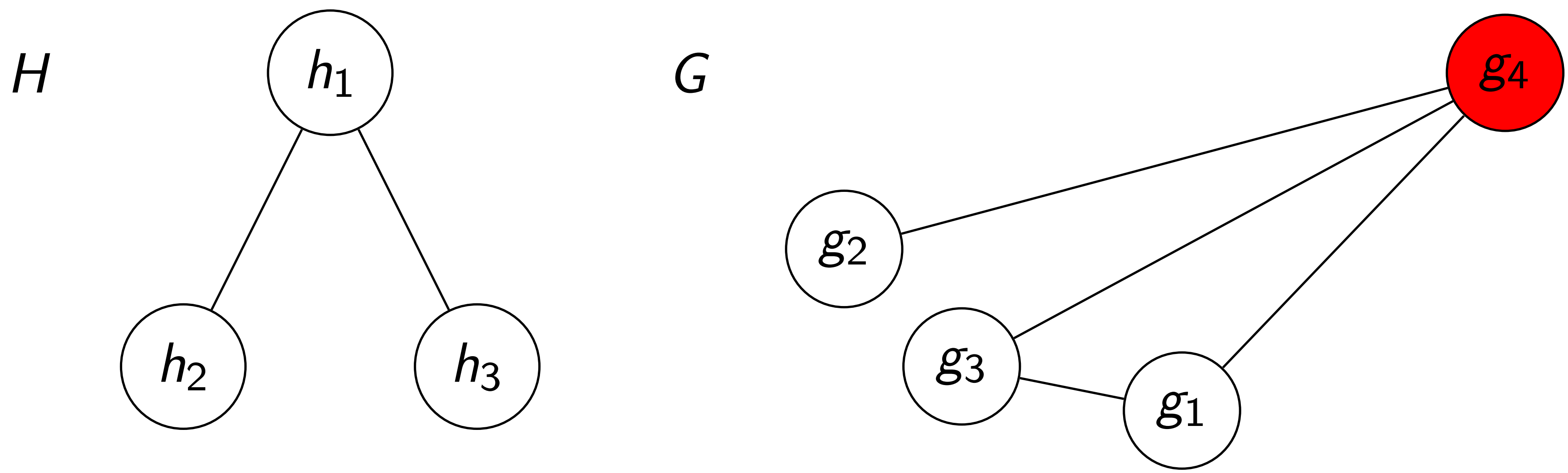
Example



Count $= 10$

Number of copies of H in $H = 2$

Example



Count $= 10$

Number of copies of H in $H = 2$

$\rightarrow$ Number of unlabelled copies of H in $G = 10/2 = 5$