Probabilistic Graph Homomorphism

Antoine Amarilli[1], **Mikaël Monet**[1,2], Pierre Senellart[2,3]

September 15th, 2017

[1]LTCI, Télécom ParisTech, Université Paris-Saclay; Paris, France

[2]Inria Paris; Paris, France

[3]École normale supérieure, PSL Research University; Paris, France

Probabilistic Graph

A directed graph $H = (V, E)$

$$\longrightarrow \quad \longrightarrow$$

Probabilistic Graph

A directed graph $H = (V, E)$

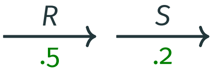
With labels $\lambda : E \to \Sigma$

Probabilistic Graph

A directed graph $H = (V, E)$

With labels $\lambda : E \to \Sigma$

With independent probability annotations $\pi : E \to [0, 1]$ on edges



Probabilistic Graph

A directed graph $H = (V, E)$

With labels $\lambda : E \to \Sigma$

With independent probability annotations $\pi : E \to [0, 1]$ on edges

This probabilistic graph represents the following probability distribution:

Probabilistic Graph

A directed graph $H = (V, E)$

With labels $\lambda : E \to \Sigma$

With independent probability annotations $\pi : E \to [0, 1]$ on edges

$$\xrightarrow[.5]{R} \; \xrightarrow[.2]{S}$$

This probabilistic graph represents the following probability distribution:

$$.5 \times .2$$

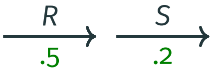
$$\xrightarrow{R} \; \xrightarrow{S}$$

Probabilistic Graph

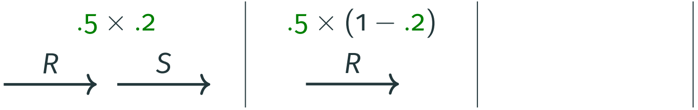
A directed graph $H = (V, E)$

With labels $\lambda : E \to \Sigma$

With independent probability annotations $\pi : E \to [0, 1]$ on edges



This probabilistic graph represents the following probability distribution:



Probabilistic Graph

A directed graph $H = (V, E)$

With labels $\lambda : E \to \Sigma$

With independent probability annotations $\pi : E \to [0, 1]$ on edges

$$\xrightarrow[.5]{R} \; \xrightarrow[.2]{S}$$

This probabilistic graph represents the following probability distribution:

$$.5 \times .2 \quad \xrightarrow{R} \; \xrightarrow{S} \quad \Big| \quad .5 \times (1 - .2) \quad \xrightarrow{R} \quad \Big| \quad (1 - .5) \times .2 \quad \xrightarrow{S} \quad \Big|$$

Probabilistic Graph

A directed graph $H = (V, E)$
With labels $\lambda : E \to \Sigma$
With independent probability annotations $\pi : E \to [0, 1]$ on edges

$$\xrightarrow[.5]{R} \; \xrightarrow[.2]{S}$$

This probabilistic graph represents the following probability distribution:

$.5 \times .2$	$.5 \times (1 - .2)$	$(1 - .5) \times .2$	$(1 - .5) \times (1 - .2)$
$\xrightarrow{R} \; \xrightarrow{S}$	$\xrightarrow{R}$	$\xrightarrow{S}$	

Graph Homomorphism

$$G = (V_G, E_G, \lambda_G) \qquad H = (V_H, E_H, \lambda_H).$$

Graph Homomorphism

$G = (V_G, E_G, \lambda_G) \qquad H = (V_H, E_H, \lambda_H)$.

$h : V_G \to V_H$ is a homomorphism iff:

Graph Homomorphism

$G = (V_G, E_G, \lambda_G) \qquad H = (V_H, E_H, \lambda_H)$.

$h : V_G \to V_H$ is a homomorphism iff:

- $(x, y) \in E_G \implies (h(x), h(y)) \in E_H$
- $(x, y) \in E_G \implies \lambda_G((x, y)) = \lambda_H((h(x), h(y)))$

Graph Homomorphism

$G = (V_G, E_G, \lambda_G) \qquad H = (V_H, E_H, \lambda_H).$

$h : V_G \to V_H$ is a **homomorphism** iff:

- $(x, y) \in E_G \implies (h(x), h(y)) \in E_H$
- $(x, y) \in E_G \implies \lambda_G((x, y)) = \lambda_H((h(x), h(y)))$

$$G = \quad x \xrightarrow{R} y \xrightarrow{S} z \xleftarrow{S} t$$
$$H = \quad \bullet \xrightarrow{R} \bullet \xrightarrow{S} \bullet \xleftarrow{R} \bullet$$

Graph Homomorphism

$G = (V_G, E_G, \lambda_G) \qquad H = (V_H, E_H, \lambda_H)$.

$h : V_G \to V_H$ is a homomorphism iff:

- $(x, y) \in E_G \implies (h(x), h(y)) \in E_H$
- $(x, y) \in E_G \implies \lambda_G((x, y)) = \lambda_H((h(x), h(y)))$

$$G = \quad \mathrm{x} \xrightarrow{R} \mathrm{y} \xrightarrow{S} \mathrm{z} \xleftarrow{S} \mathrm{t}$$

$$H = \quad \bullet \xrightarrow{R} \bullet \xrightarrow{S} \bullet \xleftarrow{R} \bullet$$

Graph Homomorphism

$G = (V_G, E_G, \lambda_G) \qquad H = (V_H, E_H, \lambda_H).$

$h : V_G \to V_H$ is a homomorphism iff:

- $(x, y) \in E_G \implies (h(x), h(y)) \in E_H$
- $(x, y) \in E_G \implies \lambda_G((x, y)) = \lambda_H((h(x), h(y)))$

$$G = \quad \mathrm{x} \xrightarrow{R} \mathrm{y} \xrightarrow{S} \mathrm{z} \xleftarrow{S} \mathrm{t}$$

$$H = \quad \bullet \xrightarrow{R} \bullet \xrightarrow{S} \bullet \xleftarrow{R} \bullet$$

Graph Homomorphism

$G = (V_G, E_G, \lambda_G) \qquad H = (V_H, E_H, \lambda_H)$.

$h : V_G \to V_H$ is a homomorphism iff:

- $(x, y) \in E_G \implies (h(x), h(y)) \in E_H$
- $(x, y) \in E_G \implies \lambda_G((x, y)) = \lambda_H((h(x), h(y)))$

$$G = \quad x \xrightarrow{R} y \xrightarrow{S} z \xleftarrow{S} t$$

$$H = \quad \bullet \xrightarrow{R} \bullet \xrightarrow{S} \bullet \xleftarrow{R} \bullet$$

Graph Homomorphism

$G = (V_G, E_G, \lambda_G) \qquad H = (V_H, E_H, \lambda_H)$.

$h : V_G \to V_H$ is a homomorphism iff:

- $(x, y) \in E_G \implies (h(x), h(y)) \in E_H$
- $(x, y) \in E_G \implies \lambda_G((x, y)) = \lambda_H((h(x), h(y)))$

$$G = \quad x \xrightarrow{R} y \xrightarrow{S} z \xleftarrow{S} t$$

$$H = \quad \bullet \xrightarrow{R} \bullet \xrightarrow{S} \bullet \xleftarrow{R} \bullet$$

Graph Homomorphism

$G = (V_G, E_G, \lambda_G) \qquad H = (V_H, E_H, \lambda_H).$

$h : V_G \to V_H$ is a homomorphism iff:

- $(x, y) \in E_G \implies (h(x), h(y)) \in E_H$
- $(x, y) \in E_G \implies \lambda_G((x, y)) = \lambda_H((h(x), h(y)))$

$$G = \quad x \xrightarrow{R} y \xrightarrow{S} z \xleftarrow{S} t$$
$$H = \quad \bullet \xrightarrow{R} \bullet \xrightarrow{S} \bullet \xleftarrow{R} \bullet$$

We write $G \rightsquigarrow H$ if there exists a homomorphism from G to H

Probabilistic Graph Homomorphism (PHom)

Let us fix:

- Finite set of labels Σ
- Class $\mathcal{G}$ of query graphs on Σ (e.g., paths, trees)
- Class $\mathcal{H}$ of instance graphs on Σ

Probabilistic Graph Homomorphism (PHom)

Let us fix:

- Finite set of labels Σ
- Class $\mathcal{G}$ of query graphs on Σ (e.g., paths, trees)
- Class $\mathcal{H}$ of instance graphs on Σ

Probabilistic Graph Homomorphism (PHom) problem for $\mathcal{G}$ and $\mathcal{H}$:

- Given a query graph $G \in \mathcal{G}$
- Given an instance graph $H \in \mathcal{H}$ and a probability valuation π
- Compute the probability that G has a homomorphism to H

Probabilistic Graph Homomorphism (PHom)

Let us fix:

- Finite set of labels Σ
- Class $\mathcal{G}$ of query graphs on Σ (e.g., paths, trees)
- Class $\mathcal{H}$ of instance graphs on Σ

Probabilistic Graph Homomorphism (PHom) problem for $\mathcal{G}$ and $\mathcal{H}$:

- Given a query graph $G \in \mathcal{G}$
- Given an instance graph $H \in \mathcal{H}$ and a probability valuation π
- Compute the probability that G has a homomorphism to H

$\rightarrow$ $\Pr(G \rightsquigarrow H) = \sum_{J \subseteq H,\ G \rightsquigarrow J} \Pr(J)$

Example

$$G = \quad x \xrightarrow{R} y \xrightarrow{S} z \xleftarrow{S} t$$

$$H = \quad \bullet \xrightarrow[.2]{R} \bullet \xrightarrow[.5]{S} \bullet \xleftarrow[.8]{R} \bullet$$

Example

$$G = \quad x \xrightarrow{R} y \xrightarrow{S} z \xleftarrow{S} t$$

$$H = \quad \bullet \xrightarrow[.2]{R} \bullet \xrightarrow[.5]{S} \bullet \xleftarrow[.8]{R} \bullet$$

$\Pr(G \rightsquigarrow H) = .2 \times .5$

Complexity of Probabilistic Graph Homomorphism

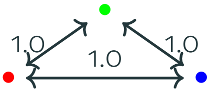
Question: what is the complexity of PHom
depending on the class $\mathcal{G}$ of query graphs and class $\mathcal{H}$ of instance graphs?

Complexity of Probabilistic Graph Homomorphism

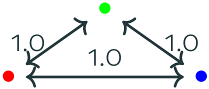
Question: what is the complexity of PHom depending on the class $\mathcal{G}$ of query graphs and class $\mathcal{H}$ of instance graphs?

Like CSP but with probabilities!

Fix one side

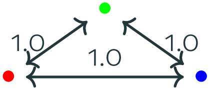
- Fix the instance graph $H =$  NP-hard

Fix one side

- Fix the instance graph $H =$ 

NP-hard
- Fix the query graph $G =$  #P-hard

Fix one side

- Fix the instance graph $H =$ 

NP-hard
- Fix the query graph $G =$  #P-hard

To make PHom tractable, we must restrict both sides

Restrict instance graphs to trees

$\mathcal{G} =$ one-way paths (1WP), $\mathcal{H} =$ polytrees (PT)

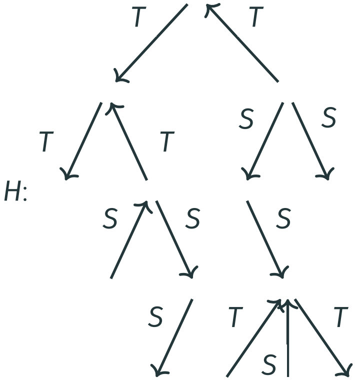
Restrict instance graphs to trees

$\mathcal{G}$ = one-way paths (1WP), $\mathcal{H}$ = polytrees (PT)

G: $\xrightarrow{T} \xrightarrow{S} \xrightarrow{S} \xrightarrow{S} \xrightarrow{T}$

Restrict instance graphs to trees

$\mathcal{G}$ = one-way paths (1WP), $\mathcal{H}$ = polytrees (PT)

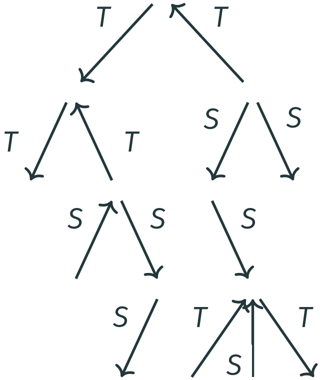


+ prob. for each edge

Restrict instance graphs to trees

$\mathcal{G}$ = one-way paths (1WP), $\mathcal{H}$ = polytrees (PT)

G: 

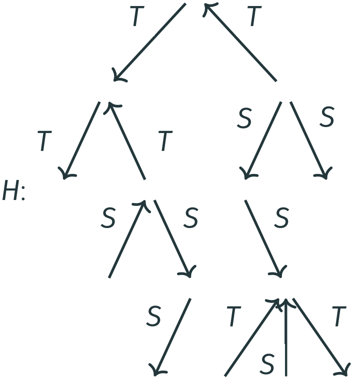
H: 

+ prob. for each edge

PHom of 1WP on PT is **#P-hard**!

$\mathcal{G}$ = *one-way paths*, $\mathcal{H}$ = *polytrees*, without labels

- What if we do not have labels?

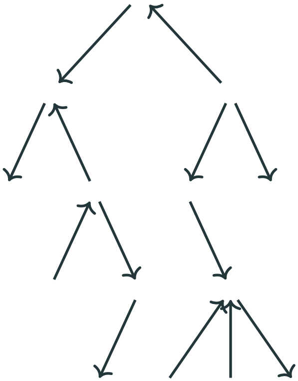


+ prob. for each edge

$\mathcal{G}$ = *one-way paths*, $\mathcal{H}$ = *polytrees*, without labels

- What if we do not have labels?

H:



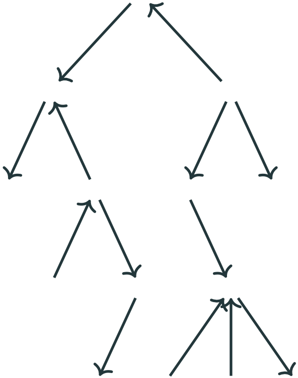
G:

+ prob. for each edge

$\mathcal{G}$ = ***one-way paths***, $\mathcal{H}$ = ***polytrees***, **without labels**

- What if we do not have labels?
- Probability that the instance graph has a path of length $|G|$

G: 

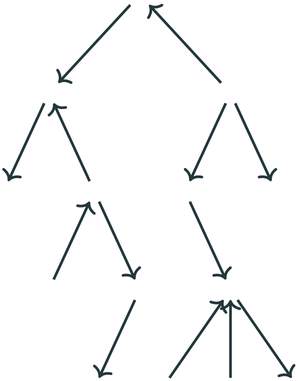
H: 

+ prob. for each edge

$\mathcal{G}$ = ***one-way paths***, $\mathcal{H}$ = ***polytrees***, **without labels**

- What if we do not have labels?
- Probability that the instance graph has a path of length $|G|$
- PTIME: Bottom-up, e.g., tree automaton

H:



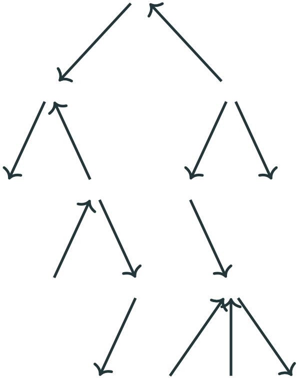
G:

+ prob. for each edge

$\mathcal{G}$ = *one-way paths*, $\mathcal{H}$ = *polytrees*, without labels

- What if we do not have labels?
- Probability that the instance graph has a path of length $|G|$
- PTIME: Bottom-up, e.g., tree automaton
- Labels have an impact!

H:



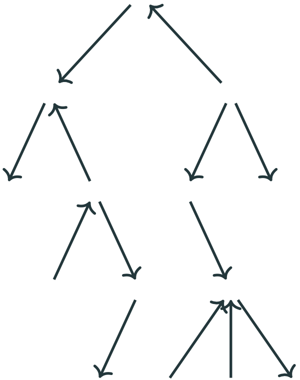
G:

+ prob. for each edge

$\mathcal{G}$ = *two-way paths*, $\mathcal{H}$ = *polytrees*, without labels

- $\mathcal{G}$ = one-way paths (1WP), $\mathcal{H}$ = polytrees (PT)

G: 

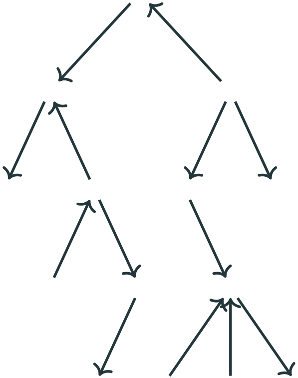
H: 

+ prob. for each edge

$\mathcal{G}$ = *two-way paths*, $\mathcal{H}$ = *polytrees*, without labels

- $\mathcal{G}$ = **two**-way paths (2WP), $\mathcal{H}$ = polytrees (PT)

H:

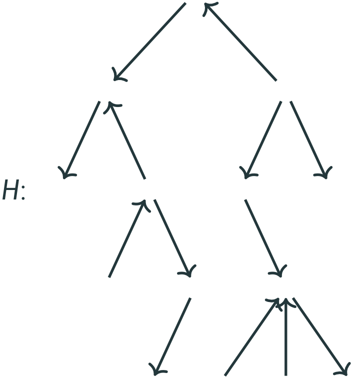


G:

+ prob. for each edge

$\mathcal{G}$ = *two-way paths*, $\mathcal{H}$ = *polytrees*, without labels

- $\mathcal{G}$ = **two**-way paths (2WP), $\mathcal{H}$ = polytrees (PT)
- **#P-hard**

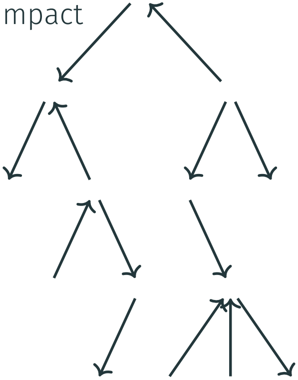


+ prob. for each edge

$\mathcal{G}$ = *two-way paths*, $\mathcal{H}$ = *polytrees*, without labels

- $\mathcal{G}$ = **two**-way paths (2WP), $\mathcal{H}$ = polytrees (PT)
- **#P-hard**
- Global orientation of the query has an impact

H:

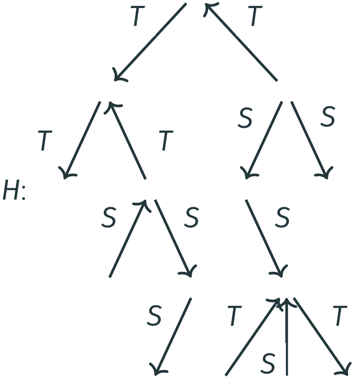


G:

+ prob. for each edge

$\mathcal{G}$ = one-way paths, $\mathcal{H}$ = downwards trees

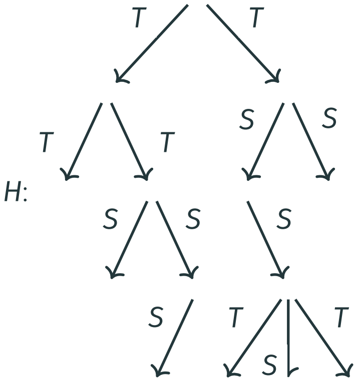
- $\mathcal{G}$ = one-way paths (1WP), $\mathcal{H}$ = polytrees (PT)



+ prob. for each edge

$\mathcal{G}$ = one-way paths, $\mathcal{H}$ = downwards trees

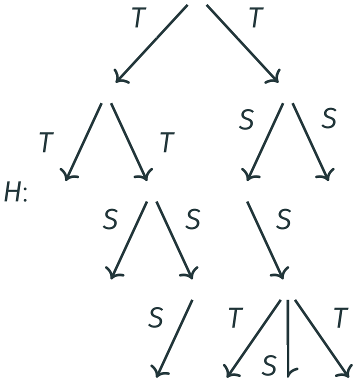
- $\mathcal{G}$ = one-way paths (1WP), $\mathcal{H}$ = **downwards trees** (DWT)



+ prob. for each edge

$\mathcal{G}$ = one-way paths, $\mathcal{H}$ = downwards trees

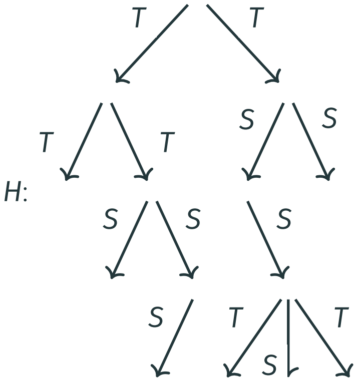
- $\mathcal{G}$ = one-way paths (1WP), $\mathcal{H}$ = **downwards trees** (DWT)
- PTIME also: *β-acyclicity of the lineage*



+ prob. for each edge

$\mathcal{G}$ = *one-way paths*, $\mathcal{H}$ = *downwards trees*

- $\mathcal{G}$ = one-way paths (1WP), $\mathcal{H}$ = **downwards trees** (DWT)
- PTIME also: *β-acyclicity of the lineage*
- Global orientation of the instance also has an impact!



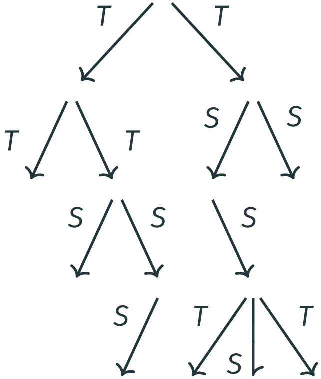
+ prob. for each edge

$\mathcal{G}$ = downwards trees, $\mathcal{H}$ = downwards trees, with labels

- $\mathcal{G}$ = one-way paths (1WP), $\mathcal{H}$ = downwards trees

G:

H:

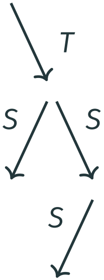


+ prob. for each edge

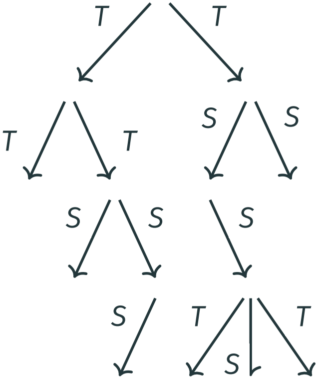
$\mathcal{G}$ = downwards trees, $\mathcal{H}$ = downwards trees, with labels

- $\mathcal{G}$ = **downwards trees** (DWT), $\mathcal{H}$ = downwards trees

G:



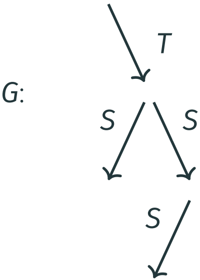
H:

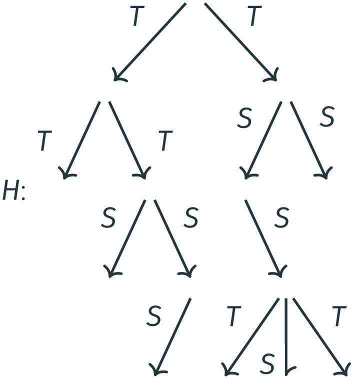


+ prob. for each edge

$\mathcal{G}$ = downwards trees, $\mathcal{H}$ = downwards trees, with labels

- $\mathcal{G}$ = **downwards trees** (DWT), $\mathcal{H}$ = downwards trees
- **#P-hard**

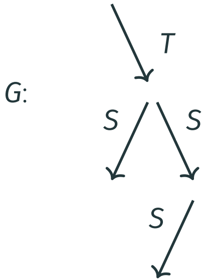


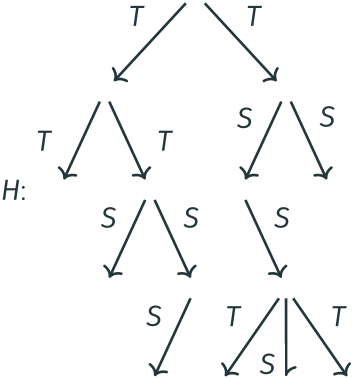


+ prob. for each edge

$\mathcal{G}$ = downwards trees, $\mathcal{H}$ = downwards trees, with labels

- $\mathcal{G}$ = **downwards trees** (DWT), $\mathcal{H}$ = downwards trees
- **#P-hard**
- Branching has an impact!





+ prob. for each edge

Results

$\downarrow G$ $\quad H\rightarrow$	1WP	2WP	DWT	PT	Connected
1WP					
2WP					
DWT		PTIME			
PT					#P-hard
Connected					

$\geqslant 2$ labels

$\downarrow G$ $\quad H\rightarrow$	1WP	2WP	DWT	PT	Connected
1WP					
2WP					
DWT		PTIME			
PT					#P-hard
Connected					

No labels

Conclusion

- We identified the complexity of PHom for a variety of graph classes

Conclusion

- We identified the complexity of PHom for a variety of graph classes
- Showed the importance of various features on the problem: **labels, global orientation, branching, connectedness**

Conclusion

- We identified the complexity of PHom for a variety of graph classes
- Showed the importance of various features on the problem: **labels, global orientation, branching, connectedness**

Future work:

- What is the hidden logic behind these tables?
- Can we get a dichotomy?

Conclusion

- We identified the complexity of PHom for a variety of graph classes
- Showed the importance of various features on the problem: **labels, global orientation, branching, connectedness**

Future work:

- What is the hidden logic behind these tables?
- Can we get a dichotomy?

Thank you for your attention!