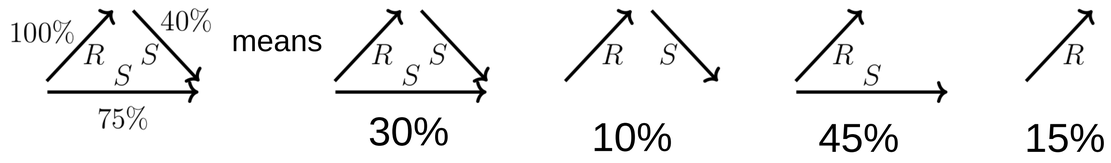
Probabilistic Graph Homomorphism: Combined Complexity

Antoine Amarilli, **Mikaël Monet**, Pierre Senellart

Probabilistic Graph Homomorphism

Probabilistic instance graph:
- Each edge is **present** or **absent** with given probability
- **Independence** across edges



Instance graph G_I — Probability distribution on graphs

Query graph G_Q: $\xrightarrow{R}\ \xrightarrow{S}\ \xleftarrow{S}$

Probabilistic graph homomorphism:
- **INPUT:** query graph G_Q and probabilistic instance graph G_I
- **OUTPUT:** probability that G_Q has a **homomorphism** to G_I

Example: for G_Q and G_I above, prob. = 30% + 10% = 40%

Known Results about Data Complexity

Data complexity:
- Fix query graph G_Q
- Study the complexity as a function of $|G_I|$

Dalvi & Suciu [1] imply:
- There is a class S of **safe** query graphs

→ Data complexity is **PTIME** if $G_Q \in S$

→ Data complexity is **#P-hard** if $G_Q \notin S$

Amarilli & al [2] imply:
- $\mathcal{G}_k$ = all graphs of **treewidth** < k

→ Data complexity of any query on $\mathcal{G}_k$ is **linear-time**

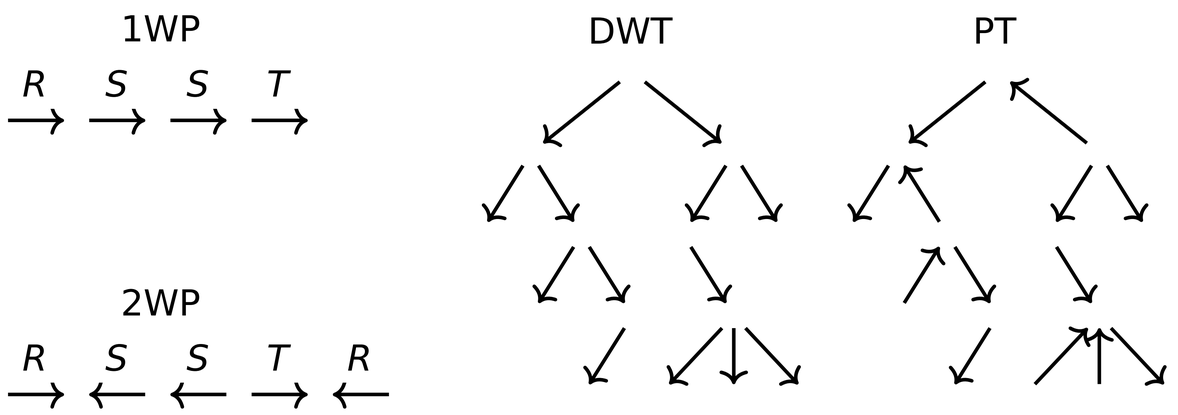
What about the **combined** complexity?
(complexity as a function of both $|G_I|$ and $|G_Q|$)

Problem

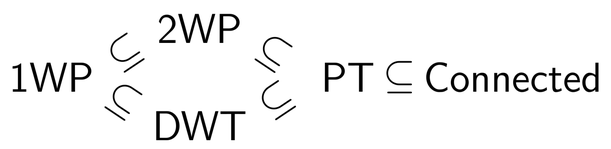
Study the **combined** complexity of the probabilistic graph homomorphism problem

Our Graph Classes

We intoduce the following graph classes: **one-way paths** (1WP), **two-way paths** (2WP), **downwards trees** (DWT) and **polytrees** (PT)



Relationship between classes



Features

- **Labeling**
- **Global orientation**
- **Branching**
- **(Connectedness)**

Results

Without labels

$\downarrow Q$ $I \rightarrow$	1WP	2WP	DWT	PT	Connected
1WP	PTIME				#P-hard
2WP					
DWT					
PT					
Connected					

With >1 labels

$\downarrow Q$ $I \rightarrow$	1WP	2WP	DWT	PT	Connected
1WP	PTIME				#P-hard
2WP					
DWT					
PT					
Connected					

Proof Techniques

- Tree automata
- β-acyclicity
- X-property
- Various coding techniques for #P-hardness (#PP2DNF and #Bipartite-Edge-Cover)

References

[1] N. Dalvi, D. Suciu
The Dichotomy of Probabilistic Inference for Unions of Conjunctive Queries
JACM, 2012

[2] A. Amarilli, P. Bourhis, P. Senellart
Provenance Circuits for Trees and Treelike Instances
Proc. ICALP, 2015

Conclusion

- First study of the combined complexity of Probabilistic Graph Homomorphism
- Shows the importance of various features
- Establishes complexity for all combinations of the graph classes we consider

However:

- Graph classes very weak
- Nowhere near a dichotomy
- Probabilistic equivalent of Feder–Vardi conjecture for combined complexity?
- Practical application? (probabilistic databases)