Learning Fast-Inference Bayesian Networks

P. R. Vaidyanathan · Stefan Szeider

December 14, 2021 · Neural Information Processing Systems

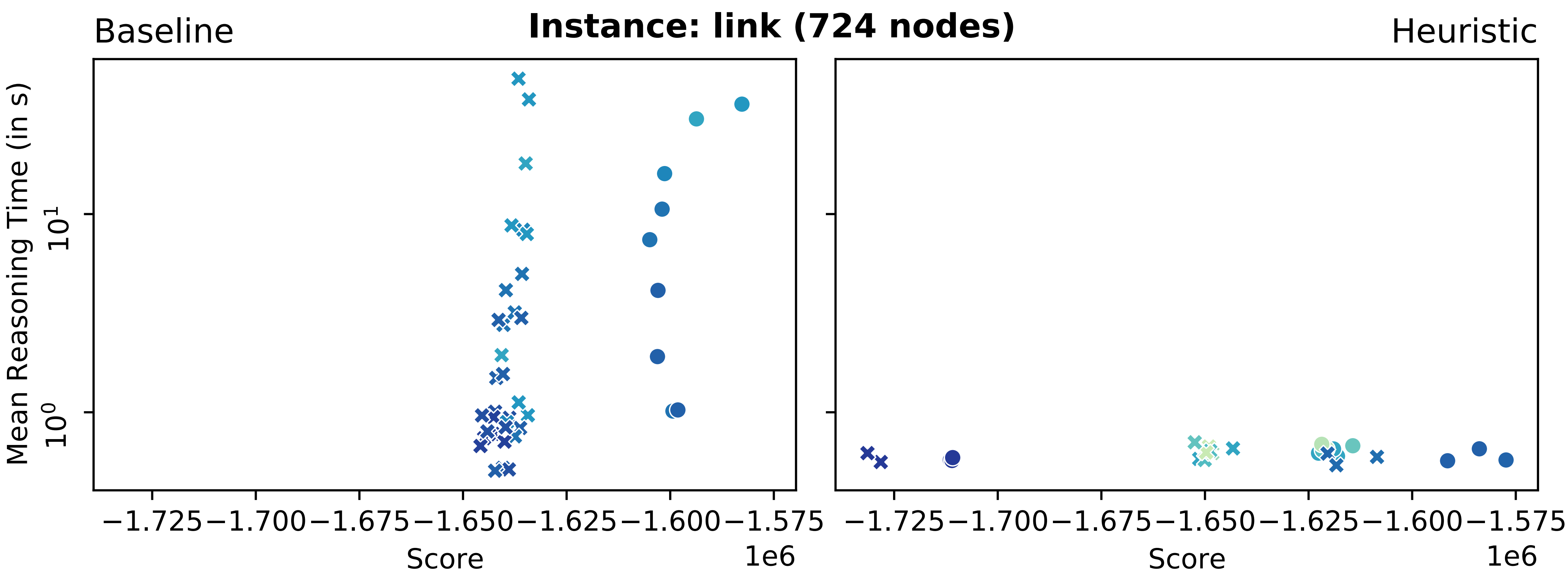
ALGORITHMS AND
COMPLEXITY GROUP

Supported by the FWF (projects P32441 and W1255) and by the WWTF (project ICT19-065)

Elevator Pitch

New approach to learn Bayesian Networks

- with *reliably low* reasoning times
- not compromising on the *quality*



Bayesian Network Basics

- A Bayesian Network (BN) is a DAG D, with random variables X as vertices along with probability distribution tables for each variable
- Arc $X_p \rightarrow X_q$ indicates q depends on p

study

	more	less
	0.7	0.3

grades

study	good	ok	bad
more	0.6	0.3	0.1
less	0.2	0.3	0.4

income

	high	low
	0.5	0.5

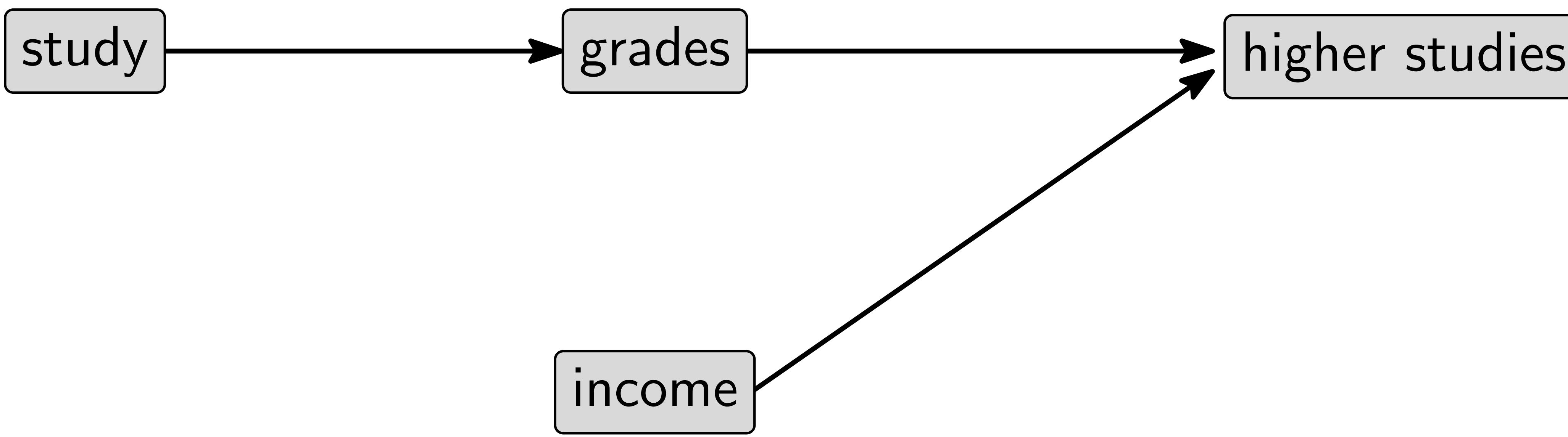
higher studies

grades/income	yes	no
good/high	0.9	0.1
good/low	0.4	0.6
ok/high	0.5	0.5
ok/low	0.3	0.7
bad/high	0.2	0.8
bad/low	0.1	0.9

Bayesian Network $\mathcal{B}$

Bayesian Network Basics

- A Bayesian Network (BN) is a DAG D, with random variables X as vertices along with probability distribution tables for each variable
- Arc $X_p \to X_q$ indicates q depends on p



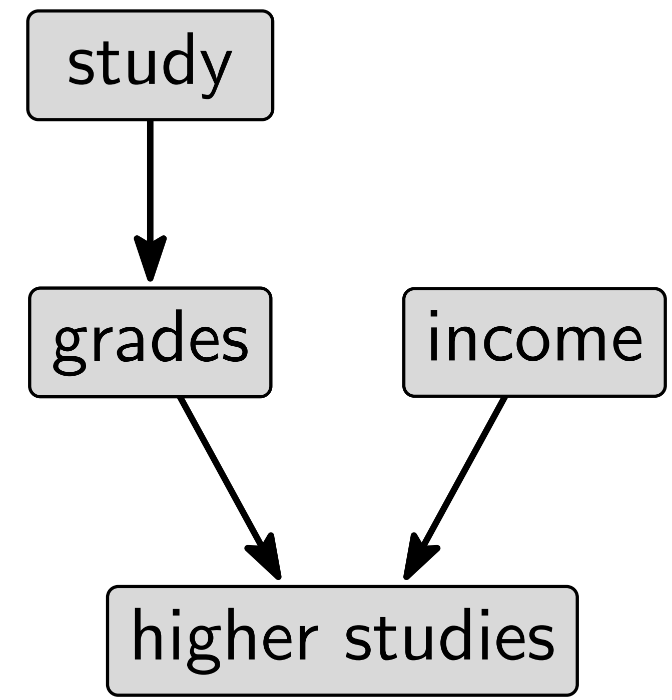
Structure of $\mathcal{B}$

Bayesian Network Basics

Bayesian Network Structure Learning (BNSL)

study	grade	income	higher studies
less	good	high	no
more	ok	low	yes
less	ok	high	yes
less	good	high	yes
⋮	⋮	⋮	⋮
more	bad	high	no
less	bad	high	no
more	bad	high	no

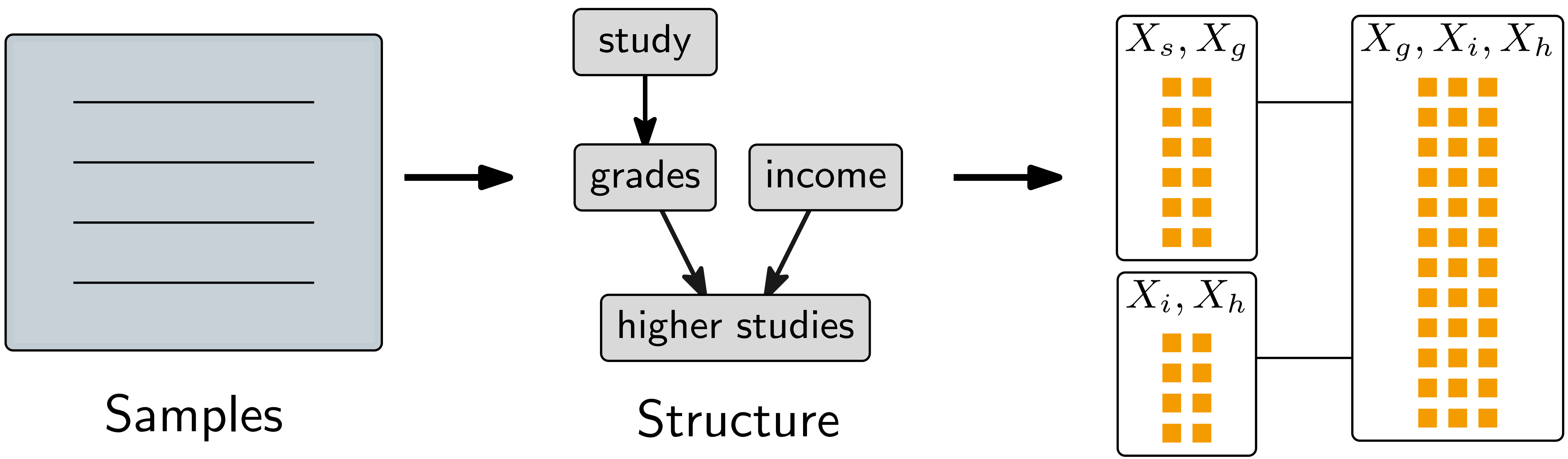
Samples



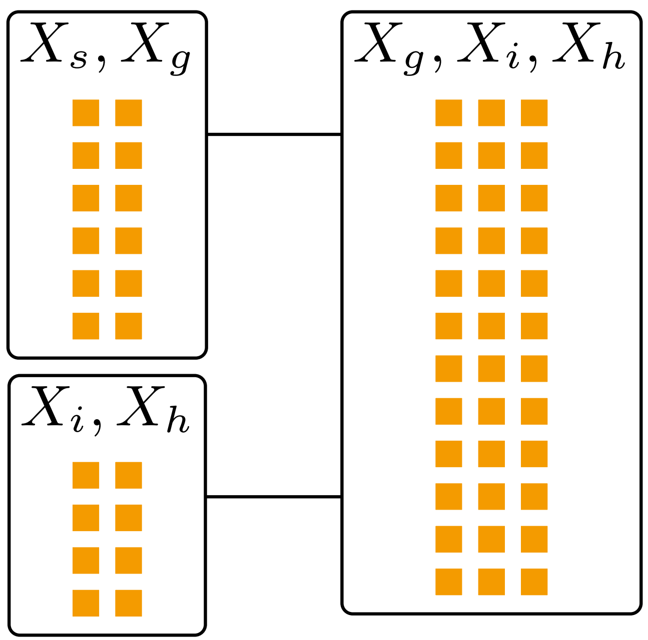
Structure

Inference on BNs

- Once learned, BNs can be used to make predictions (inference)
- Inference is **exponential** in general
- For special classes of graphs, inference can be polytime
- Most popular special class: graphs with bounded **treewidth**

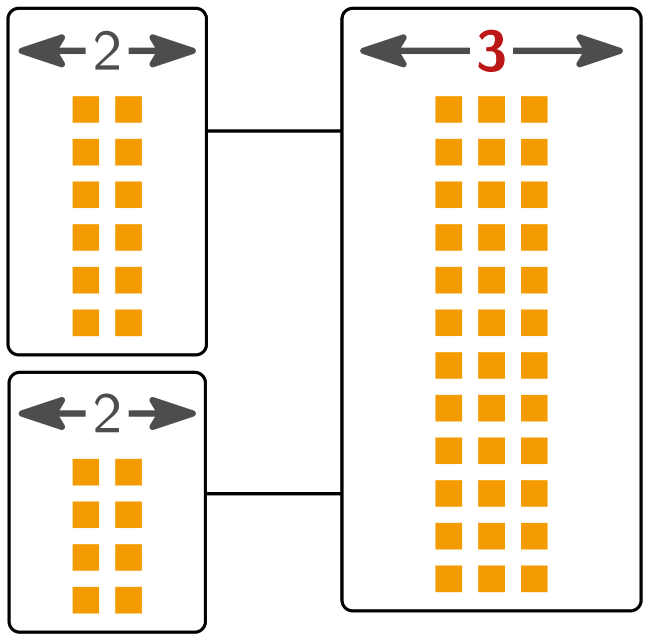


Inference on BNs



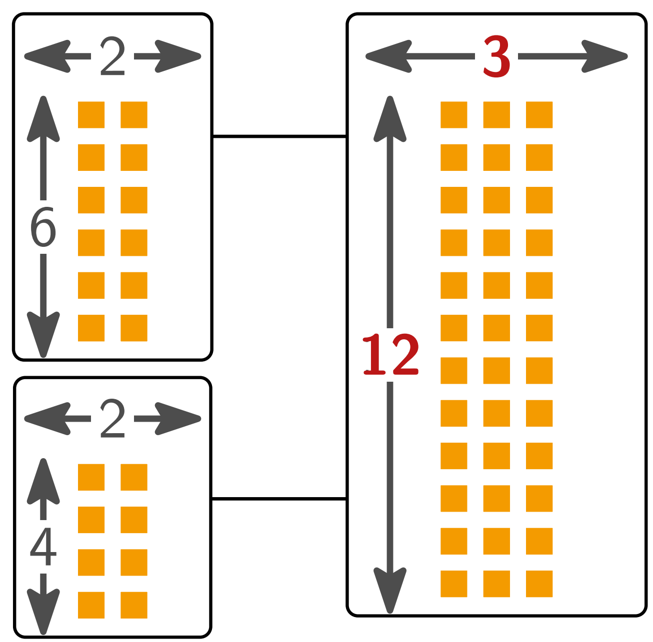
Inference on BNs

- Treewidth bounds maximum number of columns

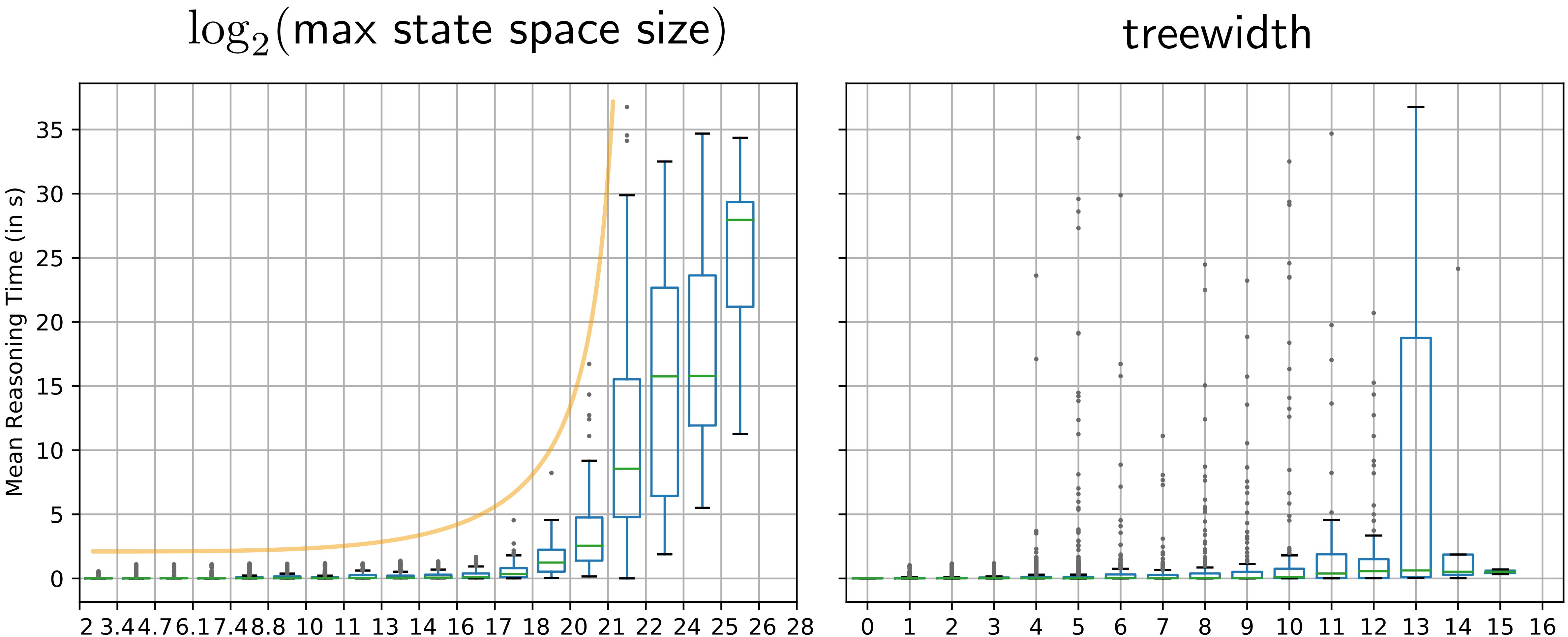


Inference on BNs

- Treewidth bounds maximum number of columns
- Many real-world datasets contain non-binary variables
- We propose to use **maximum state space size**(msss) in place of treewidth
- More fine-grained than treewidth
- Takes into account domain sizes of variables



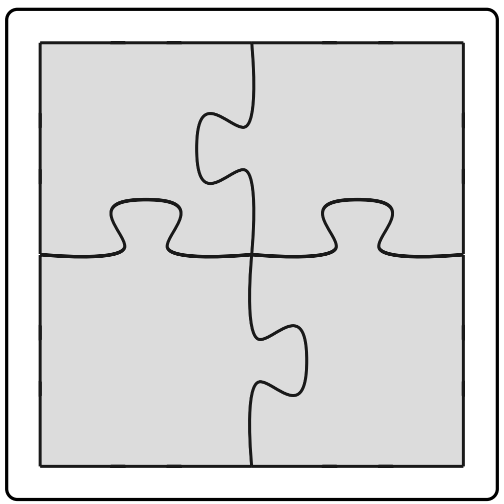
Correlation with Reasoning times



tighter correlation $\implies$ more control over reasoning time

Learning such Networks

- How to learn networks with bounded max state space size?
- Build on top of ‘SLIM’ approach used for learning bounded treewidth BNs

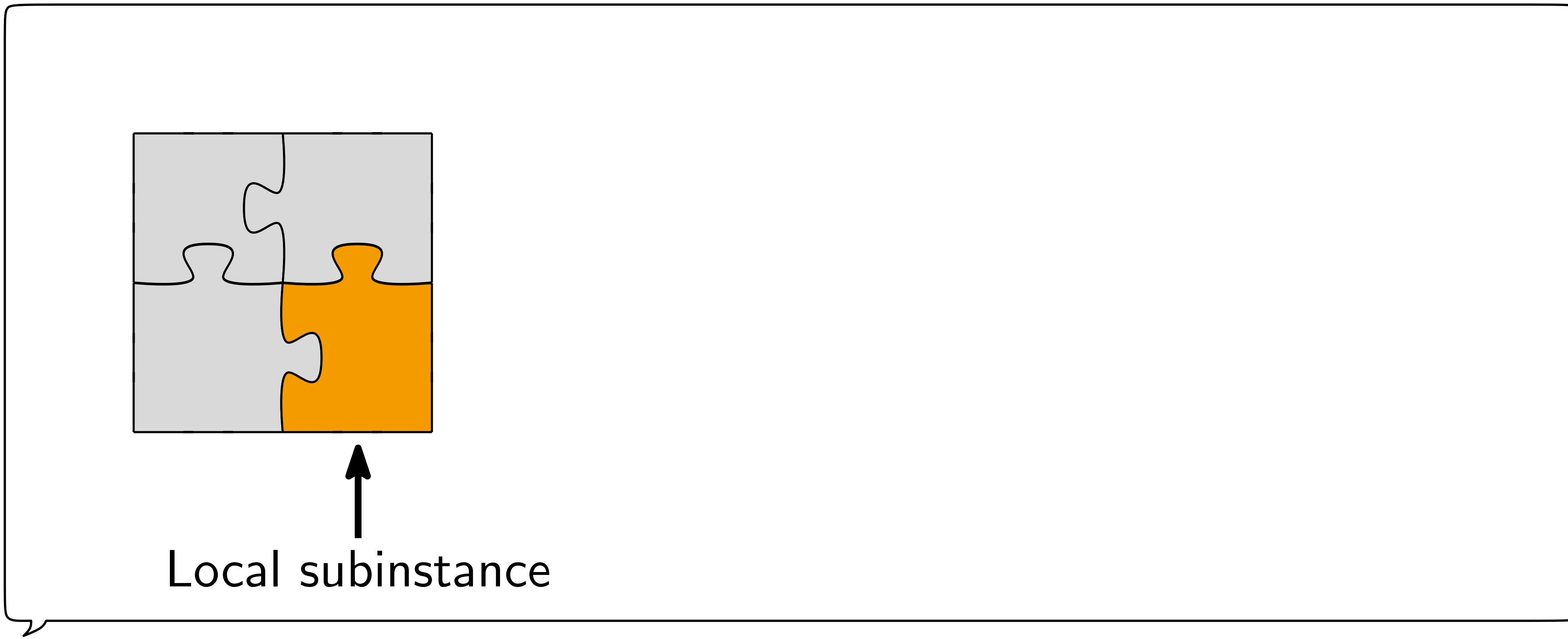


Global solution

SAT-based Local Improvement Method (SLIM)

Learning such Networks

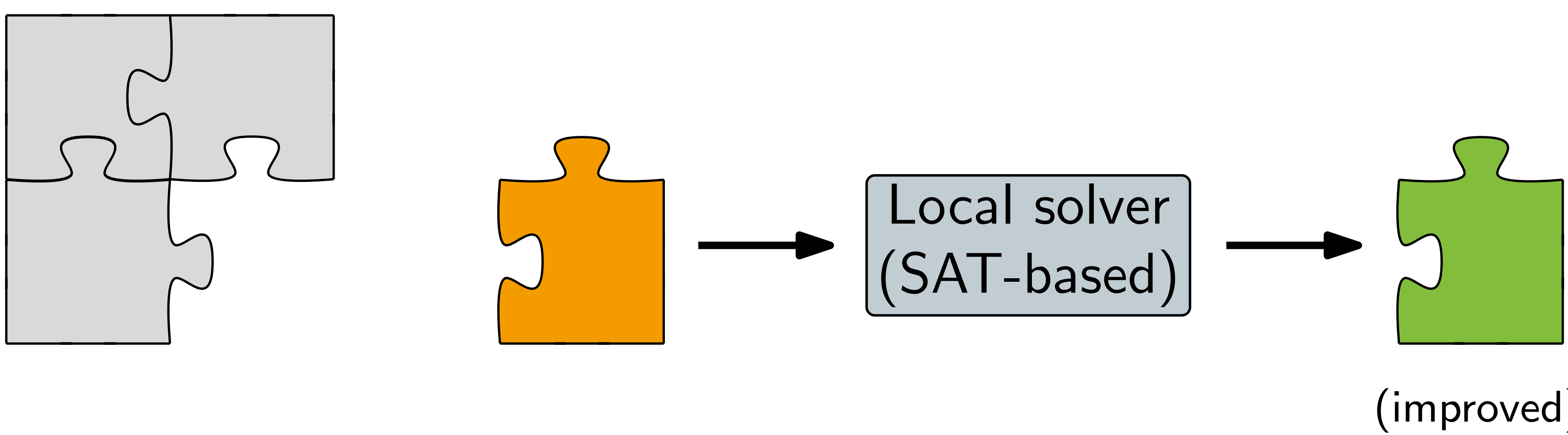
- How to learn networks with bounded max state space size?
- Build on top of 'SLIM' approach used for learning bounded treewidth BNs



SAT-based Local Improvement Method (SLIM)

Learning such Networks

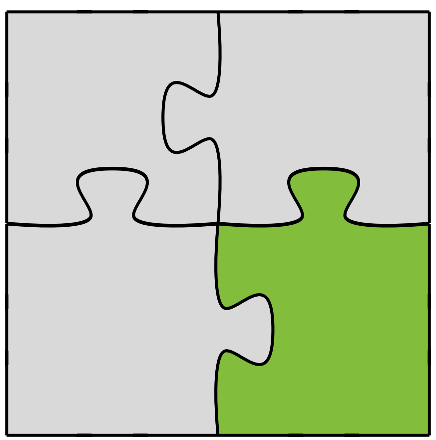
- How to learn networks with bounded max state space size?
- Build on top of 'SLIM' approach used for learning bounded treewidth BNs



SAT-based Local Improvement Method (SLIM)

Learning such Networks

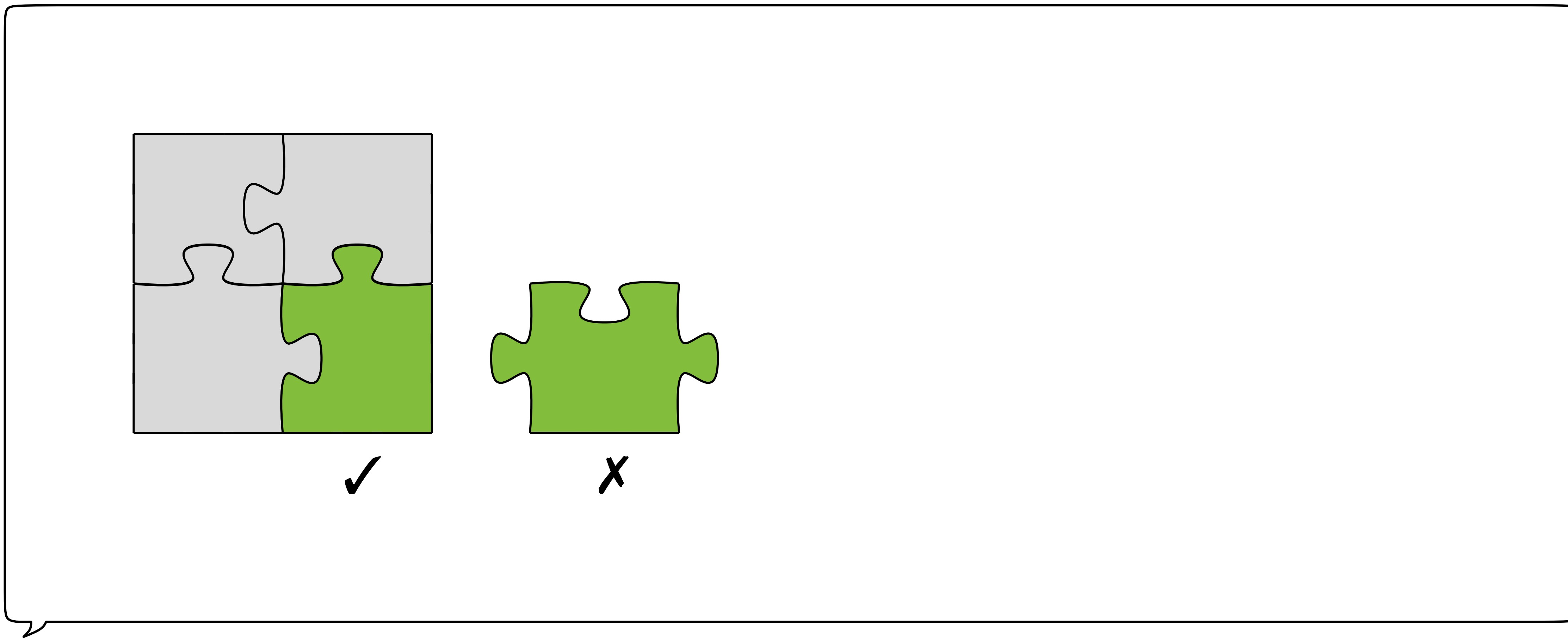
- How to learn networks with bounded max state space size?
- Build on top of 'SLIM' approach used for learning bounded treewidth BNs



SAT-based Local Improvement Method (SLIM)

Learning such Networks

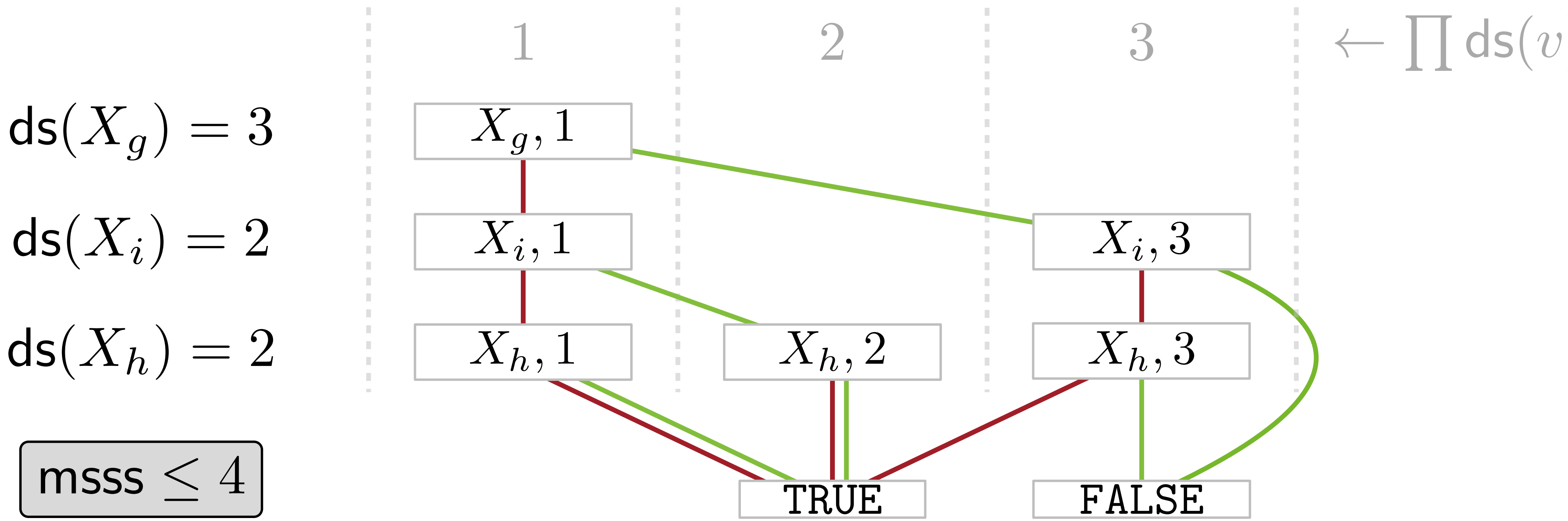
- How to learn networks with bounded max state space size?
- Build on top of 'SLIM' approach used for learning bounded treewidth BNs



SAT-based Local Improvement Method (SLIM)

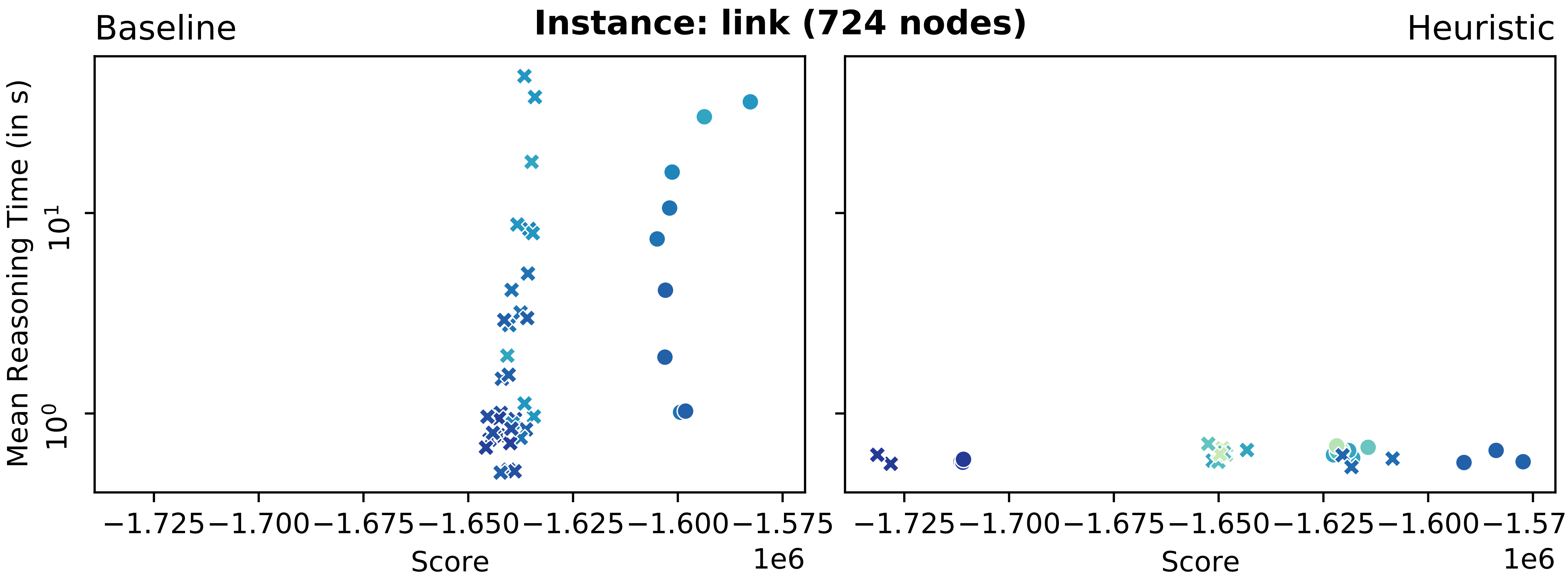
Adapting SLIM to bounded state space

- Simple cardinality counter for treewidth
- Replaced by BDD-based counter for tracking state space size



Experiments

- Modified existing heuristics to bound max state space size
- We compared several methods like original heuristic + SLIM, modified heuristic, modified heuristic + SLIM against the original heuristic as baseline
- Benchmark dataset from `bnlearn` repository, with 6-1041 random variables



Summary

- Max state space size much better indicator of inference time
- SLIM can be adapted to learn such networks (using BDDs)
- Experiments confirm inference times more reliable

Thank you!

Questions welcome