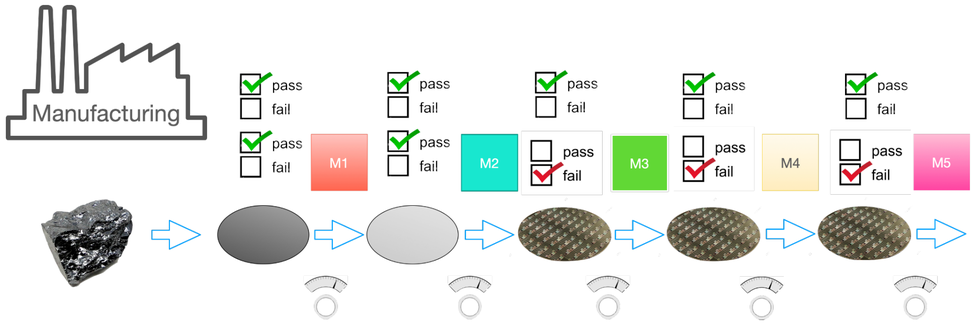
Root Cause Analysis of Failures in Microservices via Bayesian Root Cause Discovery

Kenneth Lee, Zihan Zhou, Murat Kocaoglu

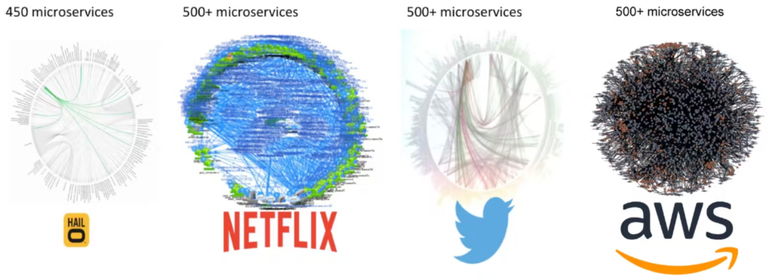
Root Cause Analysis - Why?

- Failure in a large-scale system can be difficult and costly to find.

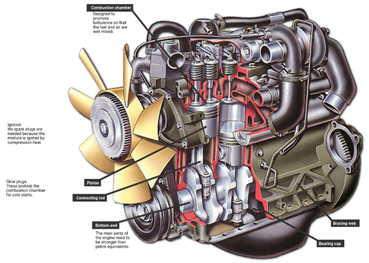
Power line fault



Tracing upstream processes for wafer defects



Latency spikes, network delay in microservices



Car Engine Failure

Processes responsible for medical errors

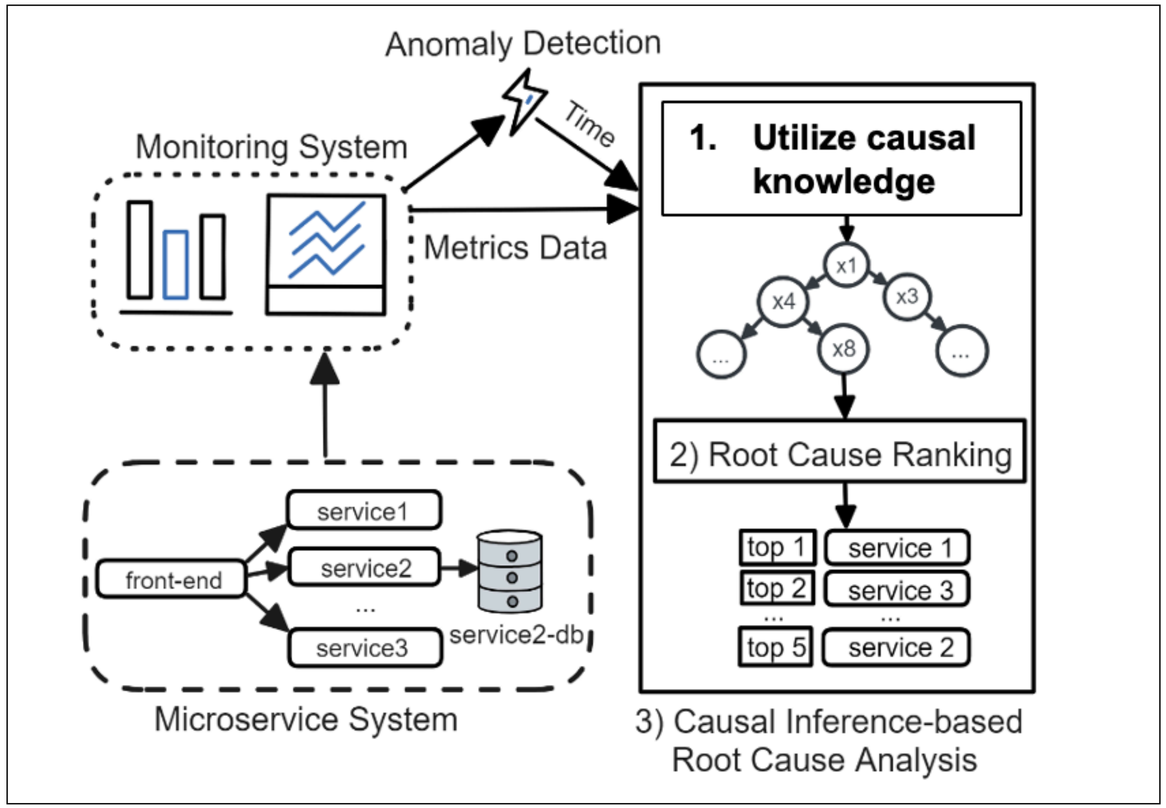
Source: https://alpacked.io/blog/microservices-use-cases/

Related Work

Non-causal methods

- Correlation-based methods
- Random walk based on a call graph
- Multi-model RCA
- LLM-based methods

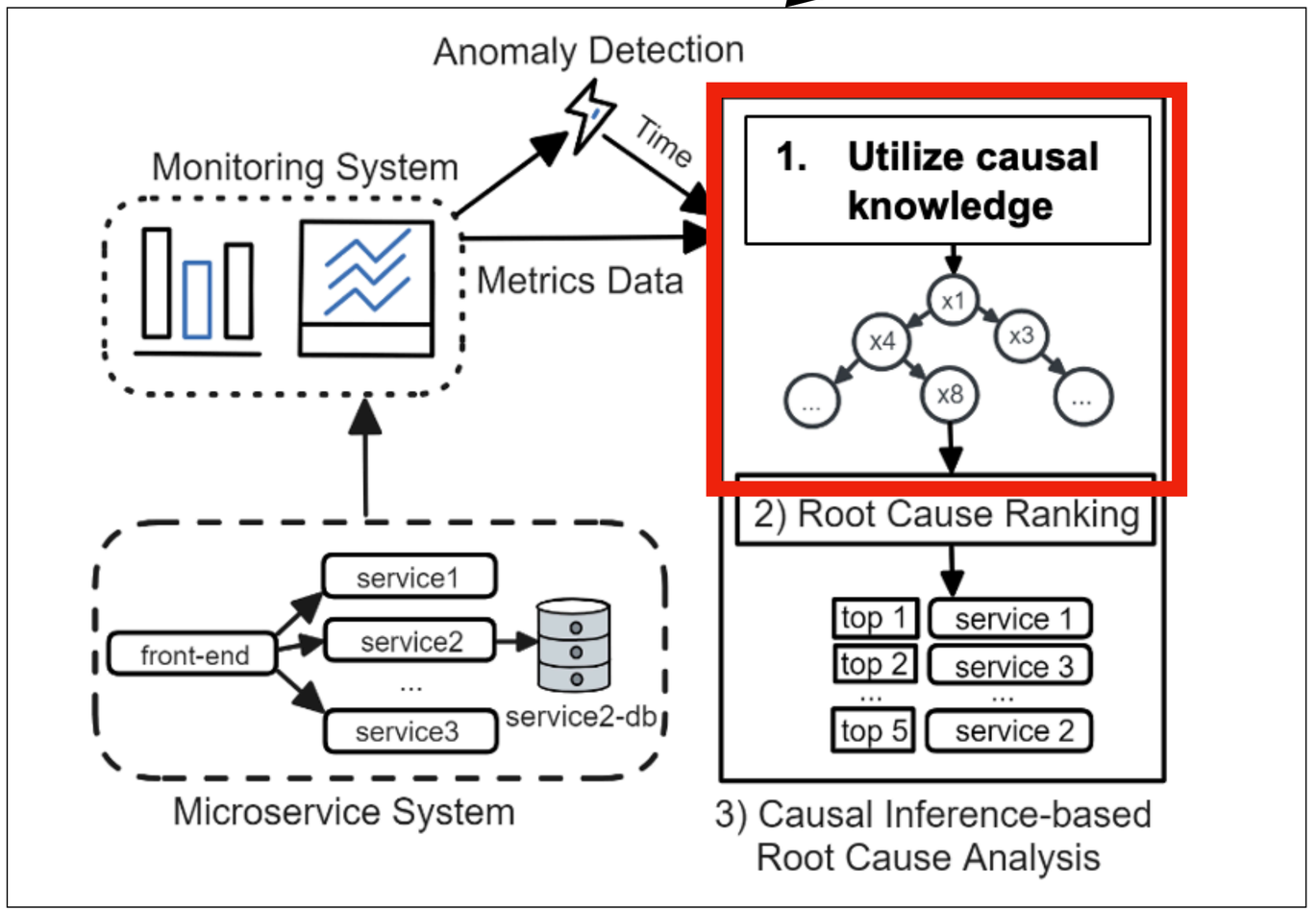
Causal methods



Pham, Luan, Huong Ha, and Hongyu Zhang. "Root cause analysis for microservice system based on causal inference: How far are we?." *Proceedings of the 39th IEEE/ACM International Conference on Automated Software Engineering*. 2024.

Contributions

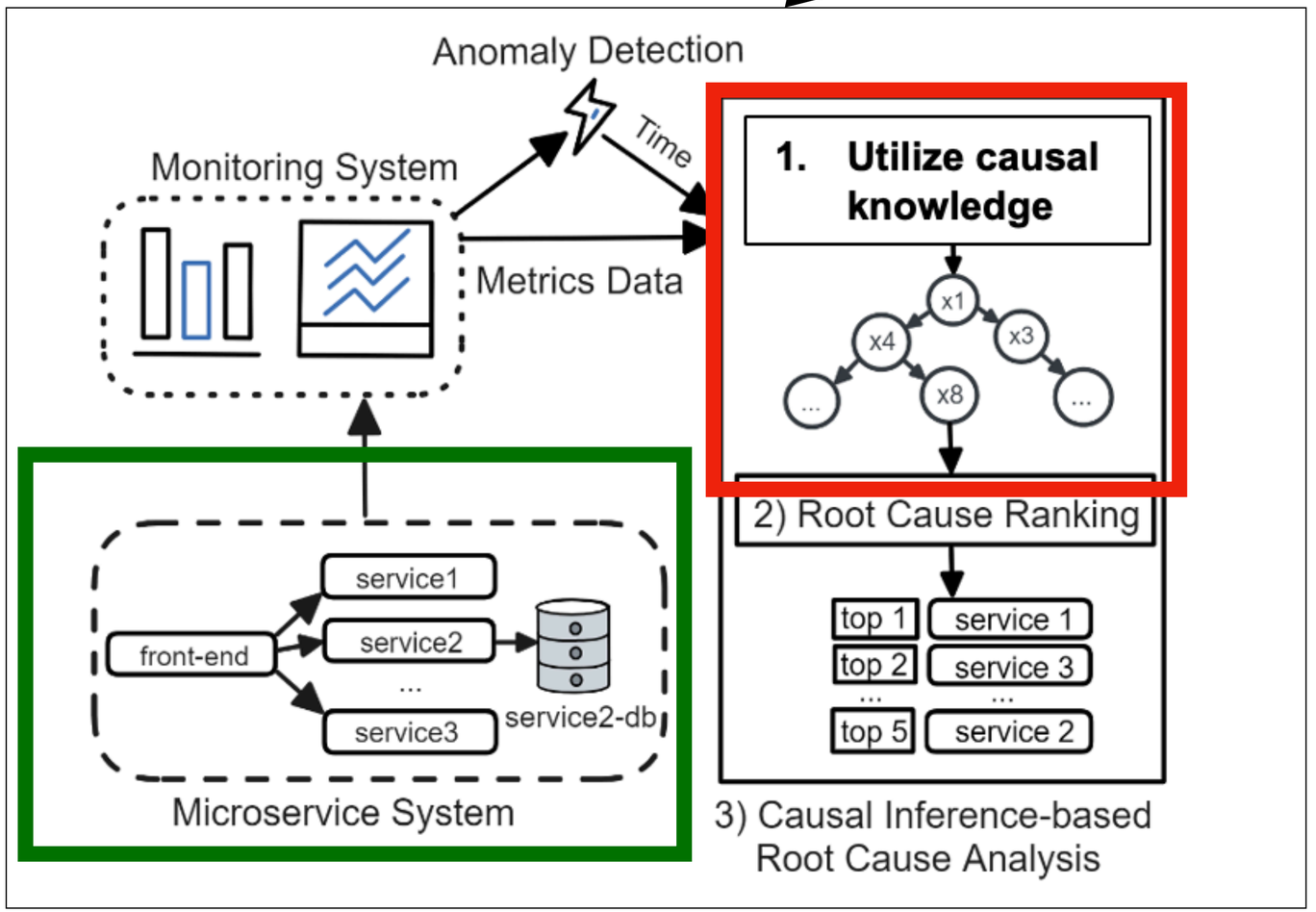
- **Problems with the existing work**
  - Do not use causal knowledge
  - Assume a full DAG available
  - Constructs a DAG only after failure occurs
  - Heuristic ranking methods (e.g. PageRank, random walk, p-value hacking)



Contributions

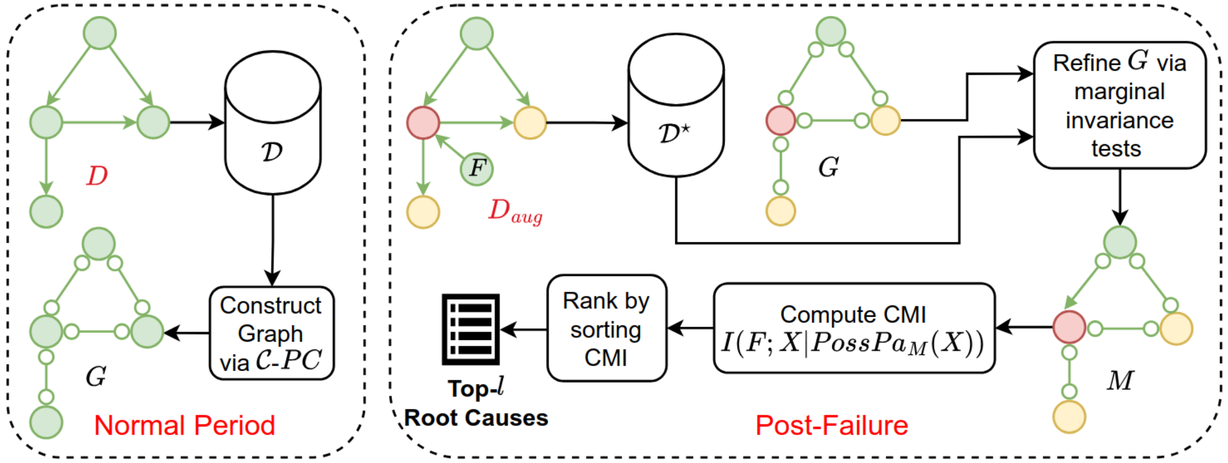
- **Problems with the existing work**
  - Do not use causal knowledge
  - Assume a full DAG available
  - Constructs a DAG only after failure occurs
  - Heuristic ranking methods (e.g. PageRank, random walk, p-value hacking)

How can we incorporate a partial causal structure learned **BEFORE** failure for RCA?



Contributions

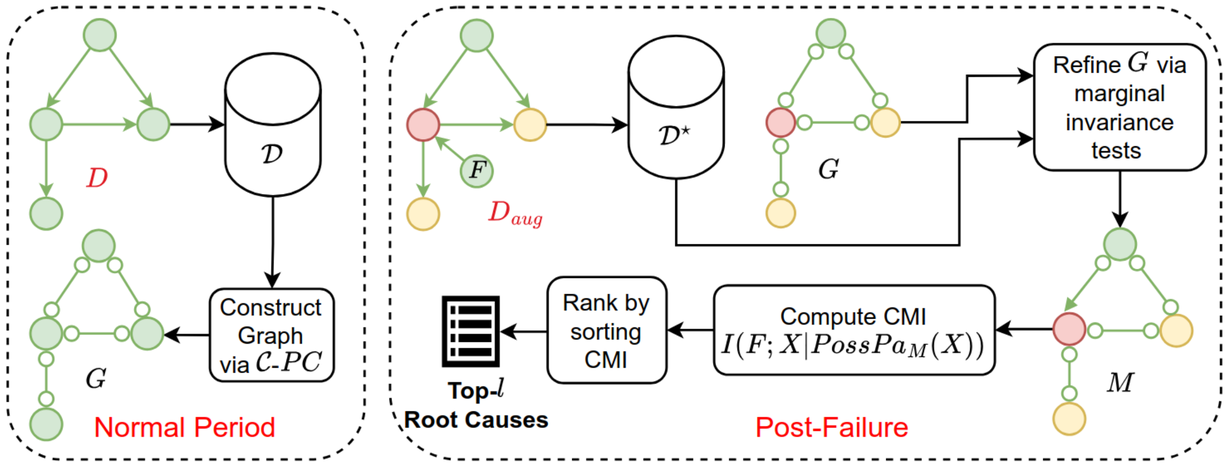
- Ikram, Azam, et al. proposed to use conditional mutual information and CI tests*
  - Failure samples **are costly** to collect, resulting in **few datapoints.**
    - **Difficult for computing conditional mutual information when the order is high**
  - **Constraint-based** methods are **prone to propagate errors** when a CI test **fails** to return correct result.



*Ikram, Azam, et al. "Root cause analysis of failures from partial causal structures." The 41st Conference on Uncertainty in Artificial Intelligence. 2025.

Contributions

- Ikram, Azam, et al. proposed to use conditional mutual information and CI tests*
  - Failure samples **are costly** to collect, resulting in **few datapoints.**
    - **Difficult for computing conditional mutual information when the order is high**
  - **Constraint-based** methods are **prone to propagate errors** when a CI test **fails** to return correct result.



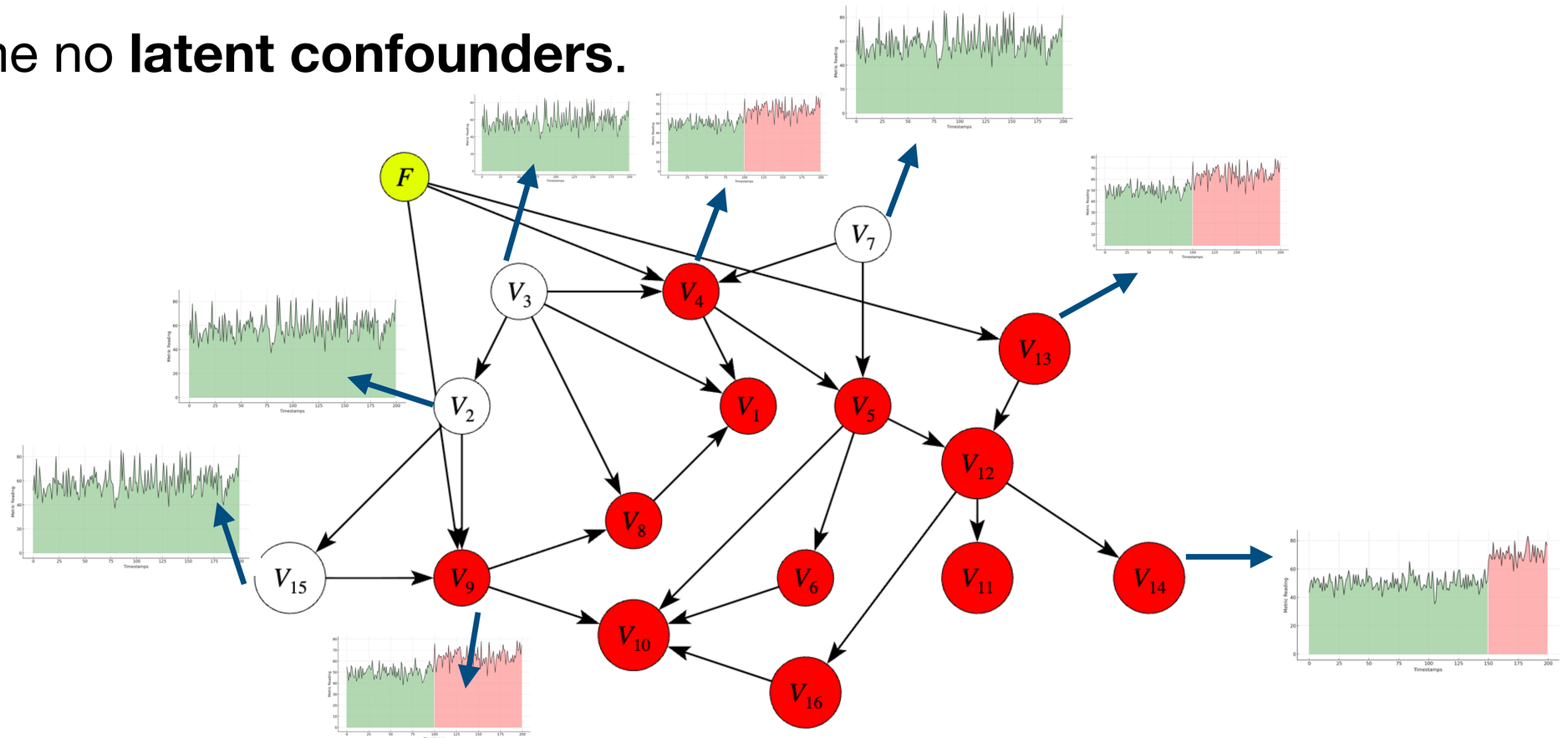
How about a Bayesian approach?

How to estimate $P(R \mid Data)$?

*Ikram, Azam, et al. "Root cause analysis of failures from partial causal structures." The 41st Conference on Uncertainty in Artificial Intelligence. 2025.

Modeling failure as interventions

- Model the failure in a system as **a soft intervention** (Ikram et al. 2022)
- Assume the **timestamp** at which the anomaly occurred is known.
- Assume no **latent confounders**.



Ikram, Azam, et al. "Root cause analysis of failures in microservices through causal discovery." *Advances in Neural Information Processing Systems* 35 (2022): 31158-31170.

Bayesian Root Cause Discovery (BRCD)

Key Idea

$$P(R \mid Data) = \frac{P(Data \mid R)P(R)}{\sum_{R'} P(Data \mid R')P(R')}$$

$$P(Data \mid R) = \sum_{G \in [G^\star]} P(Data \mid G, R)P(G \mid R)$$

(Markov equivalence class where ground truth $G^\star$ is)

Bayesian Root Cause Discovery (BRCD)

Key Idea

$$P(R \mid Data) = \frac{P(Data \mid R)P(R)}{\sum_{R'} P(Data \mid R')P(R')}$$

$$P(Data \mid R) = \sum_{G \in [G^\star]} P(Data \mid G, R)P(G \mid R)$$

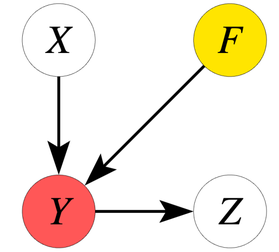
(Markov equivalence class where ground truth $G^\star$ is)

Wienöbst et al. 2023 shows that sampling a DAG from MEC and counting the size of MEC only takes polynomial-time

Wienöbst, Marcel, Max Bannach, and Maciej Liśkiewicz. "Polynomial-time algorithms for counting and sampling markov equivalent dags with applications." Journal of Machine Learning Research 24.213 (2023): 1-45.

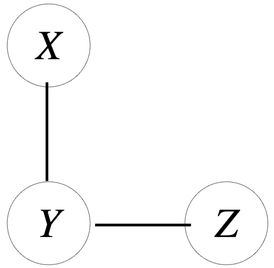
Bayesian Root Cause Discovery (BRCD)

Key Idea - an example



Ground Truth $G^{\star}_{aug}$

Assume a CPDAG is given

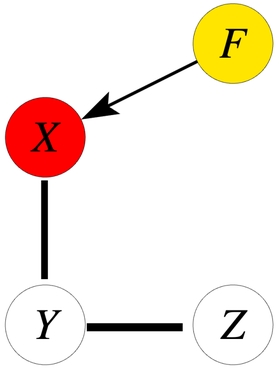


CPDAG

$$P(R \mid Data) = \frac{P(Data \mid R)P(R)}{\sum_{R'} P(Data \mid R')P(R')}$$

$$P(Data \mid R) = \sum_{G \in [G^{\star}]} P(Data \mid G, R)P(G \mid R)$$

1. Let $R = X$

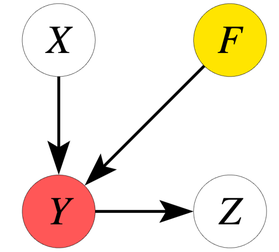


Wienöbst, Marcel, Max Bannach, and Maciej Liśkiewicz. "Polynomial-time algorithms for counting and sampling markov equivalent dags with applications." Journal of Machine Learning Research 24.213 (2023): 1-45.

Meek, C. "Causal Inference and Causal Explanation with Background Knowledge." Proc. Conf. on Uncertainty in Artificial Intelligence (UAI-95). 1995.

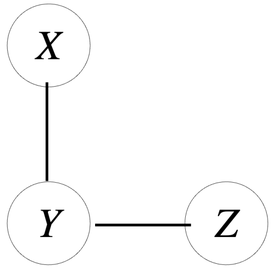
Bayesian Root Cause Discovery (BRCD)

Key Idea - an example



Ground Truth $G^{\star}_{aug}$

Assume a CPDAG is given



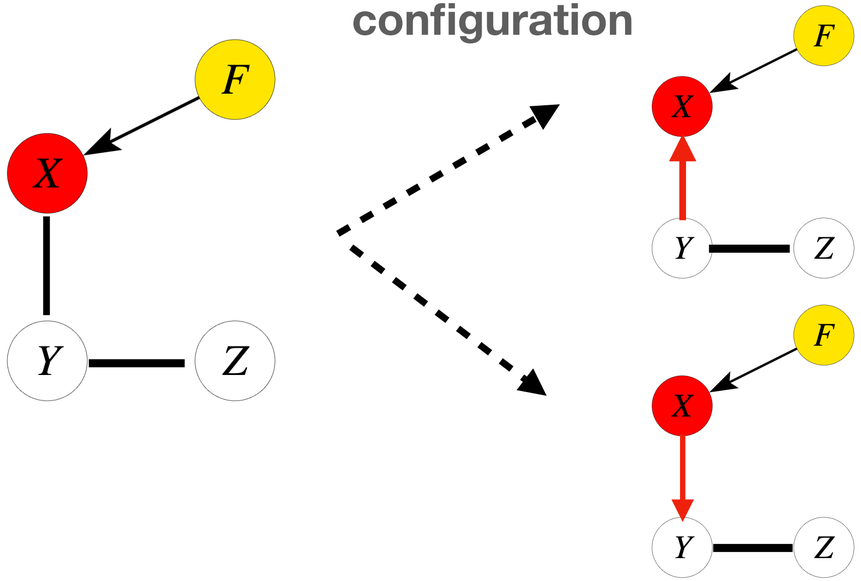
CPDAG

$$P(R \mid Data) = \frac{P(Data \mid R)P(R)}{\sum_{R'} P(Data \mid R')P(R')}$$

$$P(Data \mid R) = \sum_{G \in [G^{\star}]} P(Data \mid G, R)P(G \mid R)$$

1. Let $R = X$

2. Sample each cut configuration

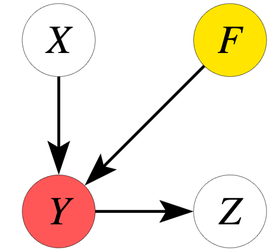


Wienöbst, Marcel, Max Bannach, and Maciej Liśkiewicz. "Polynomial-time algorithms for counting and sampling markov equivalent dags with applications." Journal of Machine Learning Research 24.213 (2023): 1-45.

Meek, C. "Causal Inference and Causal Explanation with Background Knowledge." Proc. Conf. on Uncertainty in Artificial Intelligence (UAI-95). 1995.

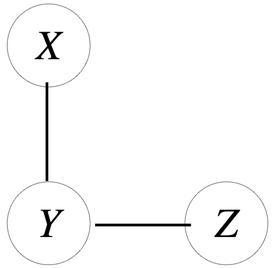
Bayesian Root Cause Discovery (BRCD)

Key Idea - an example



Ground Truth $G^{\star}_{aug}$

Assume a CPDAG is given

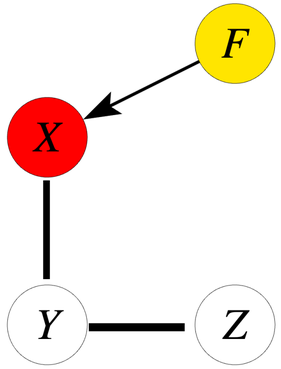


CPDAG

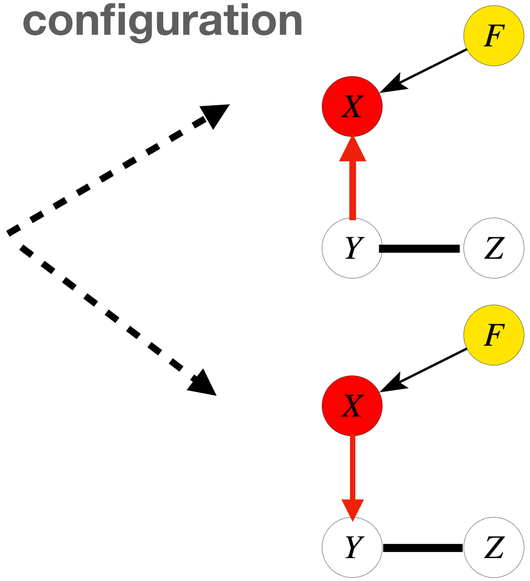
$$P(R \mid Data) = \frac{P(Data \mid R)P(R)}{\sum_{R'} P(Data \mid R')P(R')}$$

$$P(Data \mid R) = \sum_{G \in [G^{\star}]} P(Data \mid G, R)P(G \mid R)$$

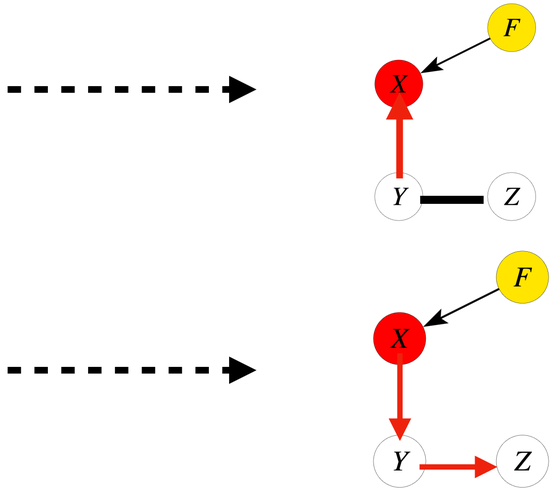
1. Let $R = X$



2. Sample each cut configuration



3. Apply Meek rules

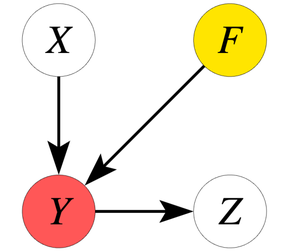


Wienöbst, Marcel, Max Bannach, and Maciej Liśkiewicz. "Polynomial-time algorithms for counting and sampling markov equivalent dags with applications." Journal of Machine Learning Research 24.213 (2023): 1-45.

Meek, C. "Causal Inference and Causal Explanation with Background Knowledge." Proc. Conf. on Uncertainty in Artificial Intelligence (UAI-95). 1995.

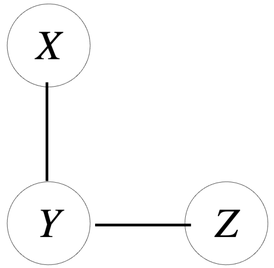
Bayesian Root Cause Discovery (BRCD)

Key Idea - an example



Ground Truth $G^{\star}_{aug}$

Assume a CPDAG is given

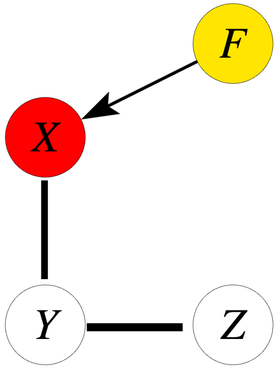


CPDAG

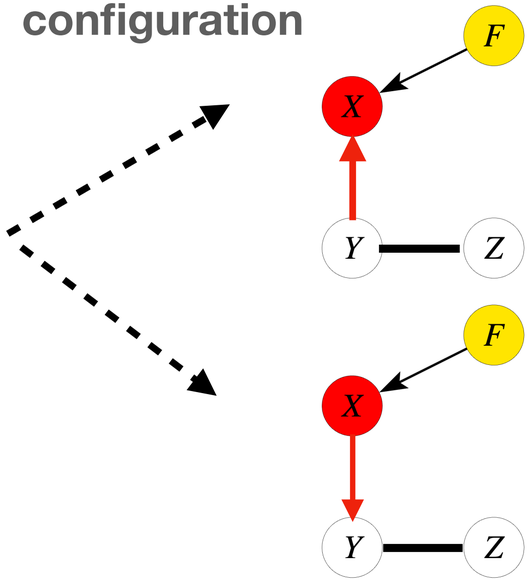
$$P(R \mid Data) = \frac{P(Data \mid R)P(R)}{\sum_{R'} P(Data \mid R')P(R')}$$

$$P(Data \mid R) = \sum_{G \in [G^{\star}]} P(Data \mid G, R)P(G \mid R)$$

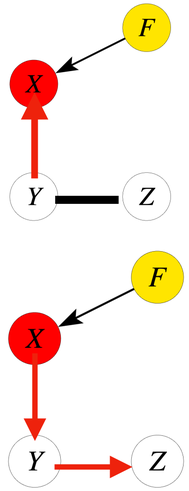
1. Let $R = X$



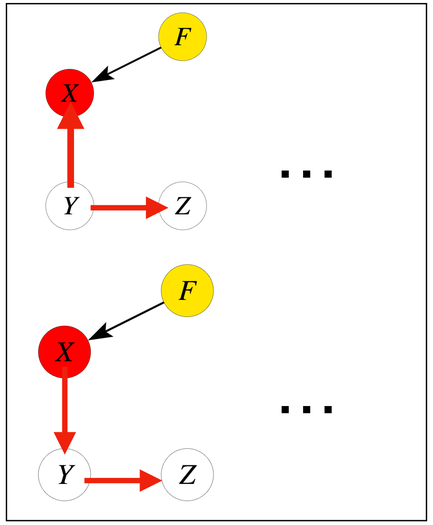
2. Sample each cut configuration



3. Apply Meek rules



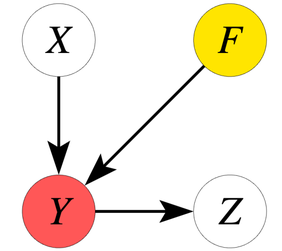
4. Sample a DAG (Wienöbst et al. 2023)



Wienöbst, Marcel, Max Bannach, and Maciej Liśkiewicz. "Polynomial-time algorithms for counting and sampling markov equivalent dags with applications." Journal of Machine Learning Research 24.213 (2023): 1-45.
Meek, C. "Causal Inference and Causal Explanation with Background Knowledge." Proc. Conf. on Uncertainty in Artificial Intelligence (UAI-95). 1995.

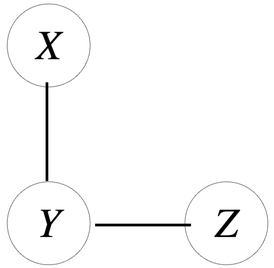
Bayesian Root Cause Discovery (BRCD)

Key Idea - an example



Ground Truth $G^{\star}_{aug}$

Assume a CPDAG is given



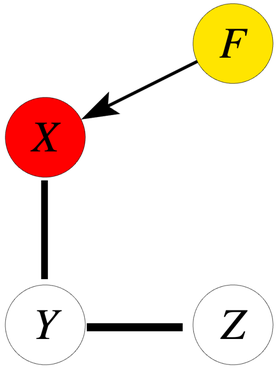
CPDAG

$$P(R \mid Data) = \frac{P(Data \mid R)P(R)}{\sum_{R'} P(Data \mid R')P(R')}$$

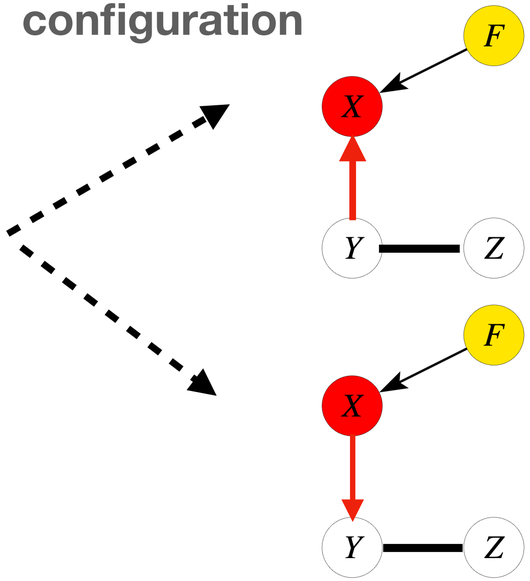
Uniform over all I-CPDAGs from each possible I-MEC

$$P(Data \mid R) = \sum_{G \in [G^{\star}]} P(Data \mid G, R)P(G \mid R)$$

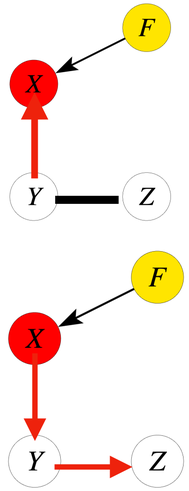
1. Let $R = X$



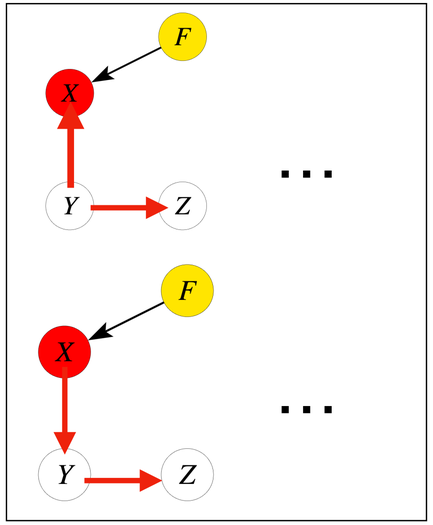
2. Sample each cut configuration



3. Apply Meek rules



4. Sample a DAG (Wienöbst et al. 2023)

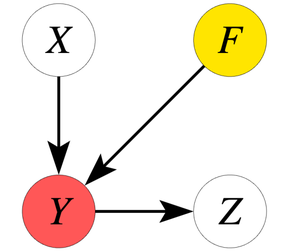


Wienöbst, Marcel, Max Bannach, and Maciej Liśkiewicz. "Polynomial-time algorithms for counting and sampling markov equivalent dags with applications." Journal of Machine Learning Research 24.213 (2023): 1-45.

Meek, C. "Causal Inference and Causal Explanation with Background Knowledge." Proc. Conf. on Uncertainty in Artificial Intelligence (UAI-95). 1995.

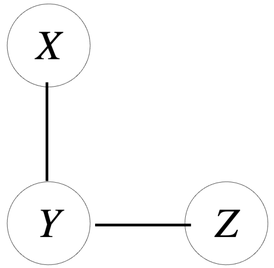
Bayesian Root Cause Discovery (BRCD)

Key Idea - an example



Ground Truth $G^{\star}_{aug}$

Assume a CPDAG is given



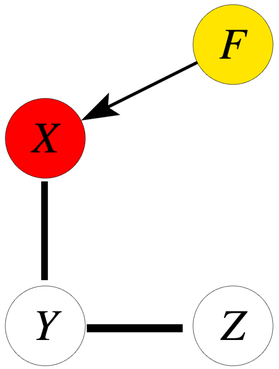
CPDAG

$$P(R \mid Data) = \frac{P(Data \mid R)P(R)}{\sum_{R'} P(Data \mid R')P(R')}$$

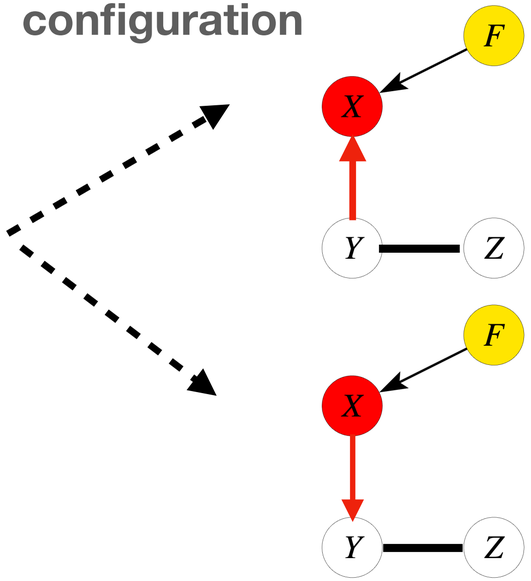
Uniform over all I-CPDAGs from each possible I-MEC

$$P(Data \mid R) = \sum_{G \in [G^{\star}]} P(Data \mid G, R)P(G \mid R)$$

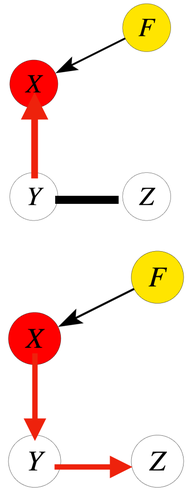
1. Let $R = X$



2. Sample each cut configuration

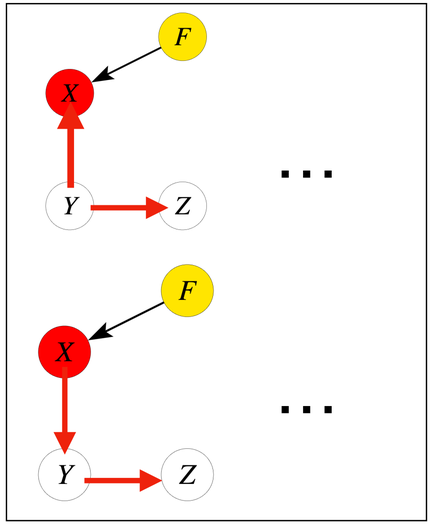


3. Apply Meek rules



4. Sample a DAG (Wienöbst et al. 2023)

(No need to enumerate all DAGs in $[G^{\star}]$)

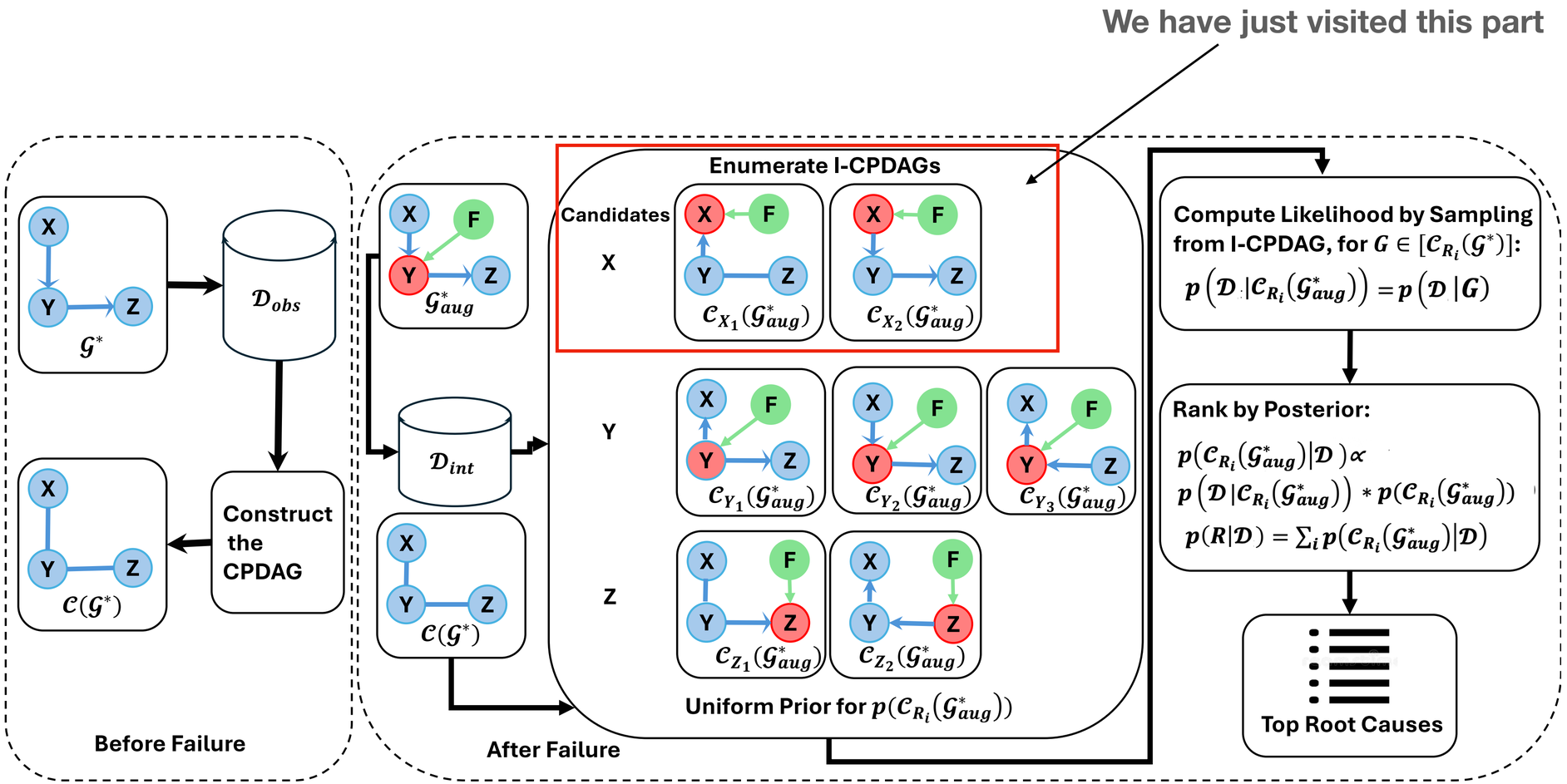


Wienöbst, Marcel, Max Bannach, and Maciej Liśkiewicz. "Polynomial-time algorithms for counting and sampling markov equivalent dags with applications." Journal of Machine Learning Research 24.213 (2023): 1-45.

Meek, C. "Causal Inference and Causal Explanation with Background Knowledge." Proc. Conf. on Uncertainty in Artificial Intelligence (UAI-95). 1995.

Bayesian RCD Algorithm

Overview



Bayesian RCD Algorithm

Robustness to Finite Sample Approximation

Assumptions

- A1 (ε-close plug in): With probability at least 1-$\eta, \eta \in (0,1)$, n is the sample size, $d_{TV}(p', p^*) \leq \varepsilon$, where $\varepsilon \to 0$ as $n \to \infty$
- A2 (Bounded log-likelihood ratios). There exists B $\leq \infty$ such that for all (G, R)

$$\left| \log \frac{p(\mathbf{X}|\mathcal{G}^\star,\mathbf{R}^\star)}{p(\mathbf{X}|\mathcal{G},\mathbf{R})} \right| \leq B.$$

Lemma (identifiability). Under modularity, positivity, for almost all parameter values, any two district sets R≠R′, if (G*,R*) is the ground truth, then

$$\Delta_{\min} := \inf_{(\mathcal{G},\mathbf{R}) \neq (\mathcal{G}^\star,\mathbf{R}^\star)} \mathbb{E}_{p^\star}\left[\log \frac{p(\mathbf{X}|\mathcal{G}^\star,\mathbf{R}^\star)}{p(\mathbf{X}|\mathcal{G},\mathbf{R})}\right] > 0.$$

Theorem (Posterior Consistency). Let R* be a set of root causes. Under A1-A2 and causal sufficiency, we have

$$p(\mathcal{G}^\star, \mathbf{R}^\star \mid \mathcal{D}) \xrightarrow[n\to\infty]{p^\star} 1.$$

Theorem (Finite Sample Bound with ε robustness). For any $\delta \in (0,1)$, let M be the number of wrong pairs (G, R) and let

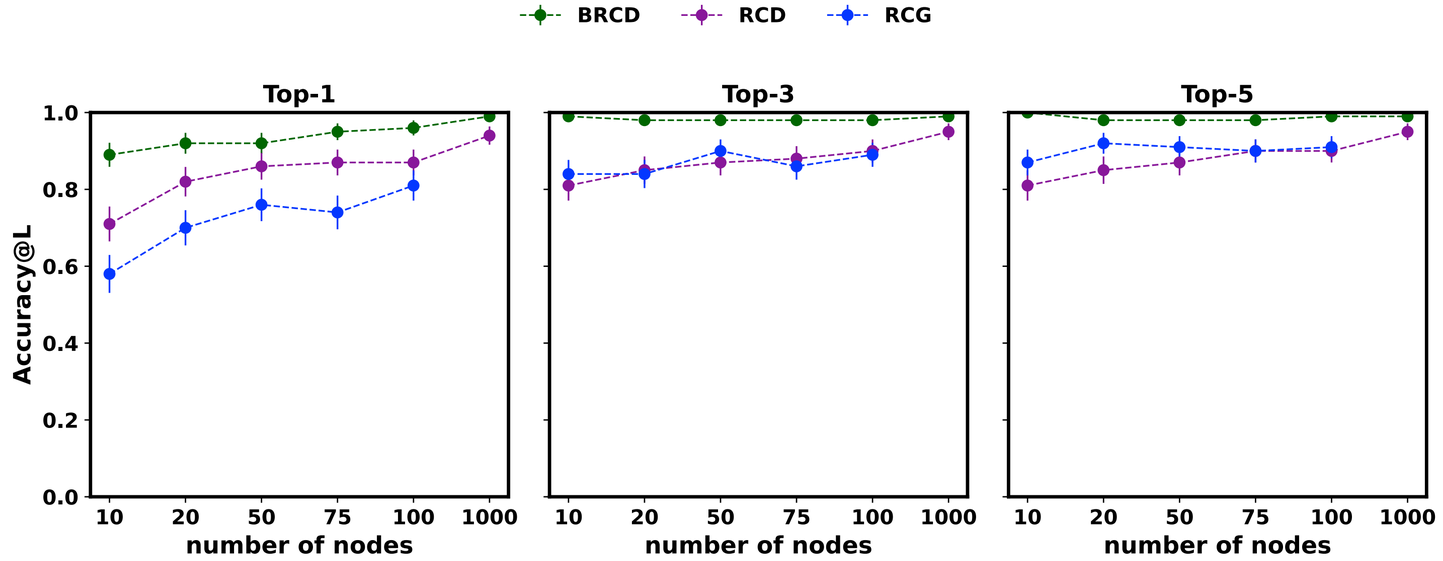
$$\Delta^{eff}_{min}(n) := \Delta_{\min} - 2B\varepsilon, t_n := B\sqrt{\frac{2\ln(2M/\delta)}{n}}$$

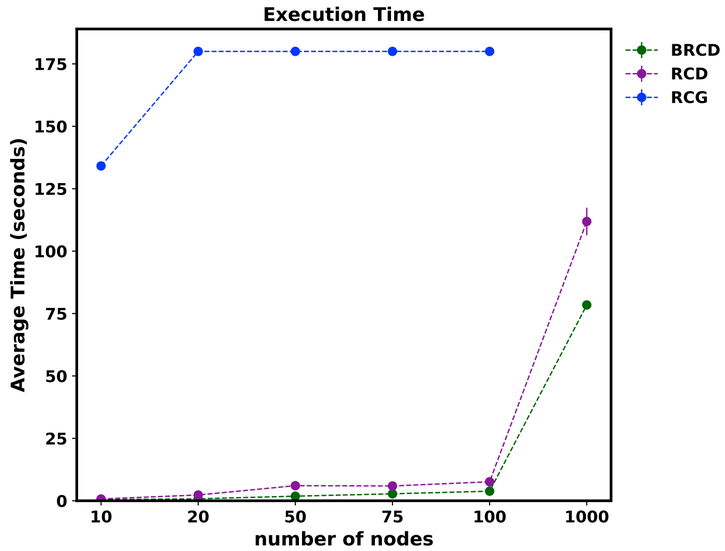
, with probability at least $1 - \delta - \eta$,

$$p(\mathcal{G}^\star, \mathbf{R}^\star \mid \mathcal{D}) \geq 1 - M\exp\left\{-n\left(\Delta^{eff}_{min}(n) - t_n\right)\right\} \max_{(\mathcal{G},\mathbf{R}) \neq (\mathcal{G}^\star,\mathbf{R}^\star)} \frac{p(\mathcal{G},\mathbf{R})}{p(\mathcal{G}^\star,\mathbf{R}^\star)}.$$

Experimental Results

Accuracy, Scalability under True CPDAG with discrete data

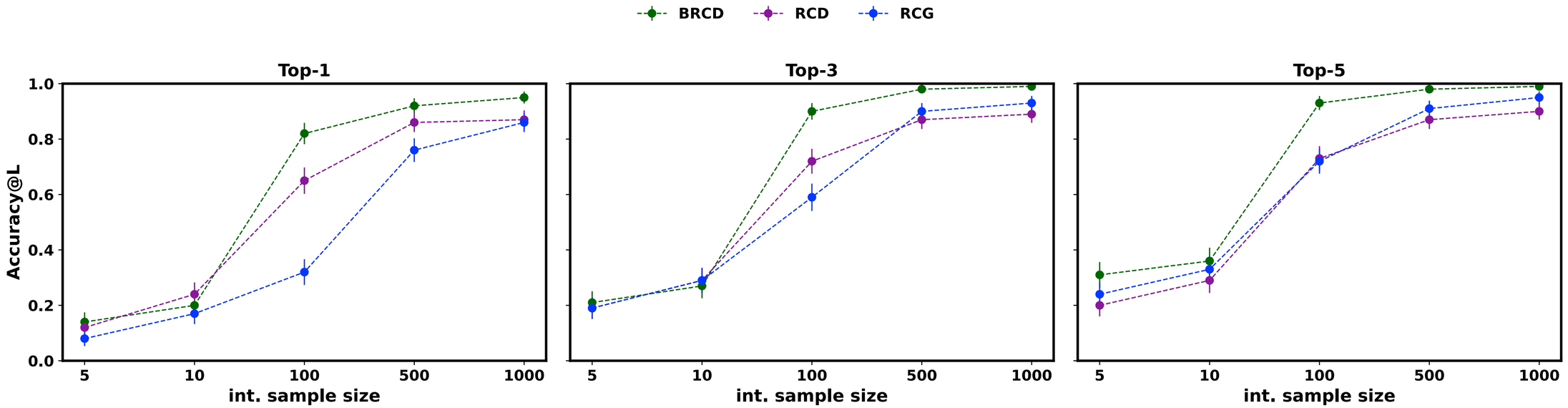




- Fixed 10k observational samples, [5, 10, 100, 500, 1000] interventional samples along x-axis; Randomly pick a single root cause
- 4 random states. Limit run time to 3 minutes
- 100 repeated experiments per node size
- True CPDAG is given to BRCD

Experimental Results

Interventional Sample Convergence



- Fixed 10k observational samples, Randomly pick a single root cause
- Fix a DAG of size 50 with 4 random states and 75 edges
- 100 repeated experiments per node size
- True CPDAG is given to BRCD

Uncertainty over CPDAGs

$$P(R \mid Data) = \frac{P(Data \mid R)P(R)}{\sum_{R'} P(Data \mid R')P(R')}$$

Bootstrapping

Instead, let C be a CPDAG rewrite as

$$P(R \mid Data) = \sum_{C} P(R \mid C, Data)P(C \mid Data)$$, where C is a $CPDAG$

$$P(C \mid Data) \propto \frac{P(Data \mid C)P(C)}{q(C)}$$

(Uniform over all sampled candidates)

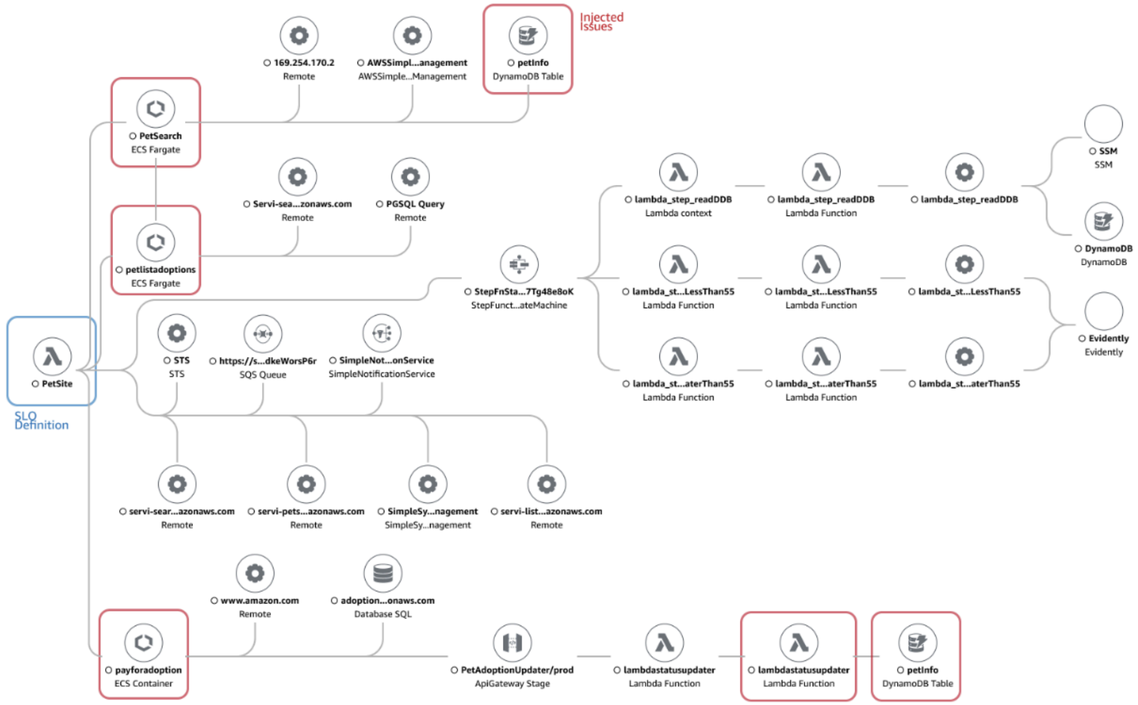
(Empirical probability From bootstrapping)

$$P(R \mid C, Data) = \frac{P(Data \mid R, C)P(R \mid C)}{\sum_{R'} P(Data \mid R', C)P(R' \mid C)}$$

$$P(Data \mid R, C) = \sum_{G \in C} P(Data \mid G, R)P(G \mid C, R)$$

Experimental Results

Real-world dataset : Petshop [1]



- Observed data sample size varies from 590 to 1656 with **5 anomalous samples only**.
- ~40 continuous variables representing the metrics of a microservice system after preprocessing
- single root cause
- Fault type: latency, service availability
- High traffic: 26 datasets, low traffic: 26 datasets, temporal traffic 1: 8 datasets, temporal traffic 2: 8 datasets
- **BRCD uses an observational discovery algorithm named BOSS to obtain a CPDAG***

[1] Hardt, Michaela, et al. "The PetShop Dataset—Finding Causes of Performance Issues across Microservices." Causal Learning and Reasoning. PMLR, 2024.
[2]Andrews, Bryan, et al. "Fast scalable and accurate discovery of dags using the best order score search and grow shrink trees." Advances in neural information processing systems 36 (2023): 63945-63956.

Experimental Results

Real-world dataset : Petshop [1]

		RCD	RCG	BARO	BRCD	BRCD-C	BRCD-M	BRCD-B10	BRCD-B100	SO	ST	Cholesky	IDI	ShapleyIQ	MicroDig	SimpleRCA
	high_traffic	0.00	0.00	0.15	0.27	**0.35**	0.08	0.31	**0.31**	0.00	0.00	0.00	0.12	0.08	0.00	0.08
	low_traffic	0.31	0.15	0.27	**0.46**	0.42	0.35	0.42	**0.46**	0.04	0.00	0.00	0.12	0.15	0.00	0.15
top-1	temporal_traffic1	0.25	0.12	0.25	0.38	**0.63**	**0.63**	0.25	0.50	0.00	0.00	0.00	0.25	0.12	0.00	0.25
	temporal_traffic2	0.62	0.25	0.25	0.50	**0.88**	**0.88**	0.62	0.62	0.00	0.00	0.00	0.25	0.25	0.00	0.25
	Average	0.30	0.13	0.23	0.40	**0.57**	0.48	0.40	0.47	0.01	0.00	0.00	0.19	0.15	0.00	0.18
	high_traffic	0.00	0.00	0.23	0.35	**0.38**	0.23	0.31	0.35	0.00	0.00	0.00	0.23	0.23	0.42	0.19
	low_traffic	0.38	0.19	0.42	0.65	0.69	0.58	**0.77**	0.73	0.04	0.04	0.00	0.23	0.38	0.62	0.31
top-3	temporal_traffic1	**0.75**	0.25	0.38	0.62	0.75	0.75	0.62	**0.75**	0.12	0.00	0.00	0.38	0.38	0.62	0.38
	temporal_traffic2	0.75	0.25	0.38	0.62	**0.88**	**0.88**	**0.88**	**0.88**	0.00	0.00	0.00	0.50	0.50	0.62	0.38
	Average	0.47	0.17	0.35	0.56	**0.68**	0.61	0.65	**0.68**	0.04	0.01	0.00	0.34	0.37	0.57	0.32
	high_traffic	0.00	0.23	0.31	0.38	**0.42**	0.38	0.35	**0.42**	0.00	0.00	0.00	0.35	0.31	0.62	0.35
	low_traffic	0.42	0.31	0.46	0.73	0.73	0.58	0.81	**0.85**	0.08	0.23	0.04	0.35	0.46	0.65	0.46
top-5	temporal_traffic1	**0.75**	0.38	0.38	0.62	**0.75**	**0.75**	0.62	**0.75**	0.25	0.00	0.00	0.38	0.38	0.62	0.38
	temporal_traffic2	0.75	0.38	0.50	0.75	**0.88**	**0.88**	**0.88**	**0.88**	0.00	0.00	0.12	0.62	0.62	0.62	0.38
	Average	0.48	0.33	0.41	0.62	0.69	0.65	0.67	**0.73**	0.08	0.06	0.04	0.43	0.44	0.63	0.39

Table 1. **Results for the Petshop dataset** (Hardt et al., 2023). Scenario-level pooled top-l accuracy of the baselines, the proposed algorithm **BRCD** and its bootstrapping variants: **BRCD-B10** and **BRCD-B100**, where **BRCD-B10** and **BRCD-B100** use 10 and 100 bootstrapped observational samples, respectively. We also include two additional **BRCD** variants: **BRCD-C** and **BRCD-M**, where **BRCD-C** directly uses the service map in Petshop data without using any CPDAG from causal discovery algorithms and **BRCD-M** enforces the ancestral relations from the service map in the causal discovery algorithm named BOSS used by **BRCD**. Each cell is a case-weighted aggregate over two cases named *availability* and *latency*. There are a total of 26, 26, 8, and 8 datasets for high_traffic, low_traffic, temporal_traffic1, and temporal_traffic2, respectively. Each dataset contains only 5 anomalous samples.

Summary

- We developed **first Bayesian root cause discovery method that leverages a polynomial-time DAG sampling method to directly identify the root cause via posterior approximation** given a partial causal structure.
- We provide the **first theoretical guarantees for nonparametric RCA**: (i) the identifiability of root causes under extended faithfulness when only a CPDAG is available (Lemma 4.1), and (ii) posterior consistency with an exponential finite-sample bound that remains valid even when the likelihood is computed using an ϵ-accurate plug-in estimator (Theorem 4.3, 4.4).
- **BRCD** outperforms the existing state-of-the-art methods in our synthetic experiments and in real-world applications such as Online Boutique, Sockshop, and Petshop in terms of overall performance