CAUSAL LEARNING USING INTERVENTION AND EXPERIMENTS

Meghamala Sinha, Stephen Ramsey, Prasad Tadepalli

School of EECS, Oregon State University

Introduction & Problem Statement

Observational data are cheap and easily avalable but can be only used to identify among Markov Equivalent structures. For associated events A & B, there is one of these explanations:

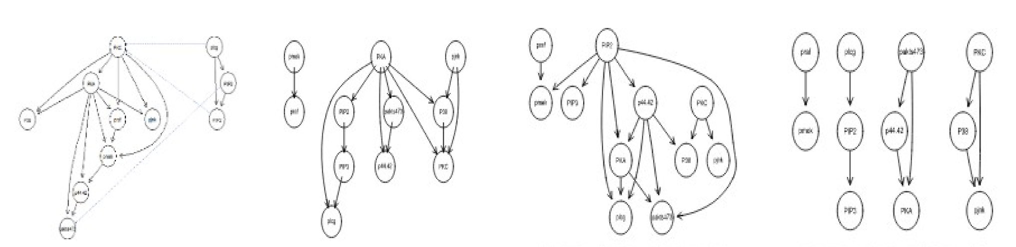
A caused B | B caused A | A and B has a common cause

❑ Why do we need to intervene?

X -> Y -> Z | X <- Y <- Z | X <- Y -> Z

- All the above 3 cases are Markov Equivalent, and encode the same conditional independence statement: X⊥Z|Y
- Intervention is needed to determine the causal structure
- Intervening on X will differentiate between direction: X->Y or Y<-X

Motivation



(a) TP: 17, FP: 0, FN: 3 (b) TP: 3, FP: 10, FN: 17 (c) TP: 5, FP: 10, FN: 15 (d) TP: 4, FP: 5, FN: 16

Figure 1 : Network inferred from (a) Sachs et. al (b) 2 observational experiments (c) Pooling data from a observational and a interventional experiment d) ``Learn and Vote'' using same experiments as (c).

❑ Effects of intervening

- Every Intervention has a specific target set
- Can interventions be "perfect"?
- Is the "graph-surgery" assumption feasible in real world?
- Are targets of intervention "deterministic"?
- Can we combine data from various such "mutilated" structures for network prediction?

Causal Network Learning

Observation of Y:
P(X, Y =1,Z) = P(Z|Y=1) * P(Y=1|X) * P(X)

Intervention on Y:
P(X, do(Y=1),Z) = P(Z|Y=1) * P(X)

❑ Assumptions

- Satisfies Causal Markov assumption
- Satisfies faithfulness assumption
- Assumes Acyclicity
- Absence of any Latent Variable
- Equal Sample size in each experiment

Methodology and Approach

❑ Popular Algorithms

- GDS
- GIES
- Simy
- ICP
- Sachs et at

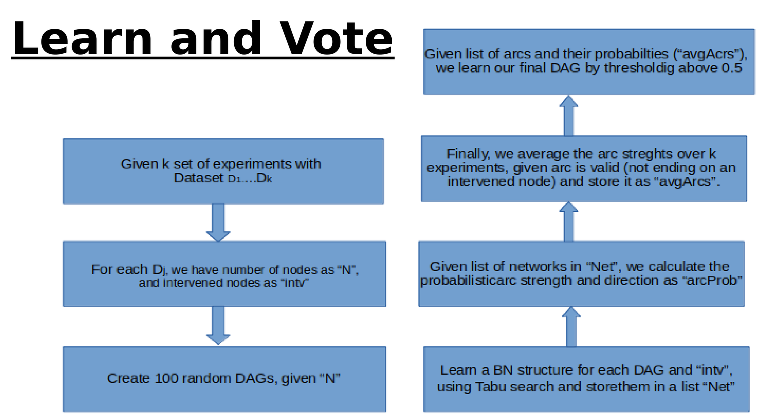
❑ Datasets

- Flow Cytometry
- Lizard
- Asia
- Alarm
- GmInt
- Insurance

Our Approach

- **Data originated after various kinds of "unknown" conditions resulted after each chemical reagent. The underlying network are no longer "structure equivalent"**
- **Sachs et al used the (Cooper & Yoo, 1999) score in their work. Implementing their method gave us an extra 8 false positive arcs along with the 17 true predictions**
- **We analyzed each experiments separately and learned a network by averaging over the results with a threshold of 0.5. This gave us significant reduction of False positives.**

❑ Learn and Vote



ALGORITHM 1 Learn and Vote

Input: set of k experiments with dataset D_1,D_2...D_k

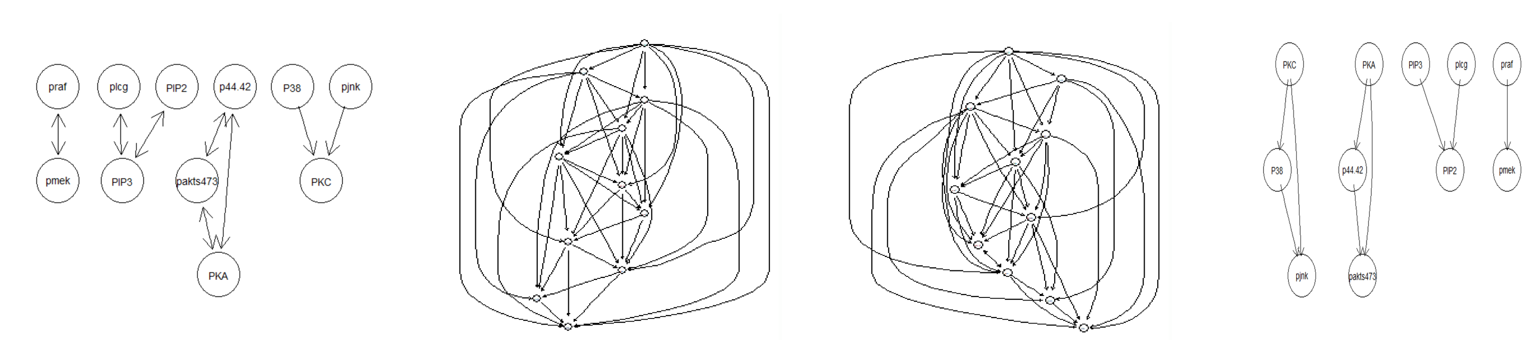
Output: DAG G^f = (E, V), final causal network

```
procedure OUR APPROACH
    for j=1 to k do
        N=nodes In D_j
        intv=Intervened nodes in D_j
        randomNet=createRandNet(N, 100)
        for l=1 to 100 do
            Net[l]=Tabu(randomNet[l], intv)
        arcProb[j]=arcStrength(Net)
    avgArcs = avgNetwork(arcProb)
    G^f = learnDAG(avgArcs,Threshold)
```

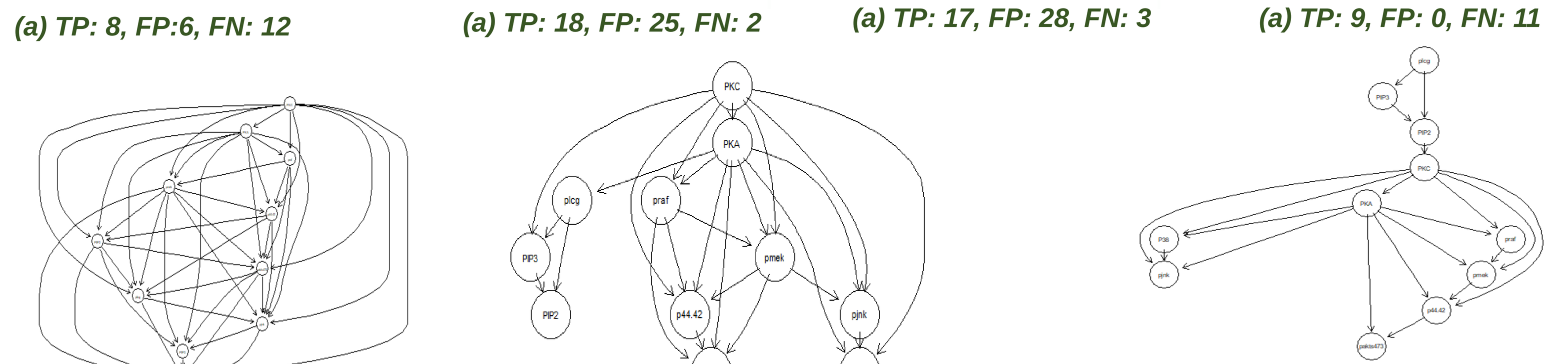
METHOD	Expected	Reported	Missed
Sachs et al.	15/17	17/17	3
Learn and Vote	16/20	20/20	2

Table1: Result on Flow-cytometry

Learned Causal Network Results



(a) TP: 8, FP:6, FN: 12 (a) TP: 18, FP: 25, FN: 2 (a) TP: 17, FP: 28, FN: 3 (a) TP: 9, FP: 0, FN: 11



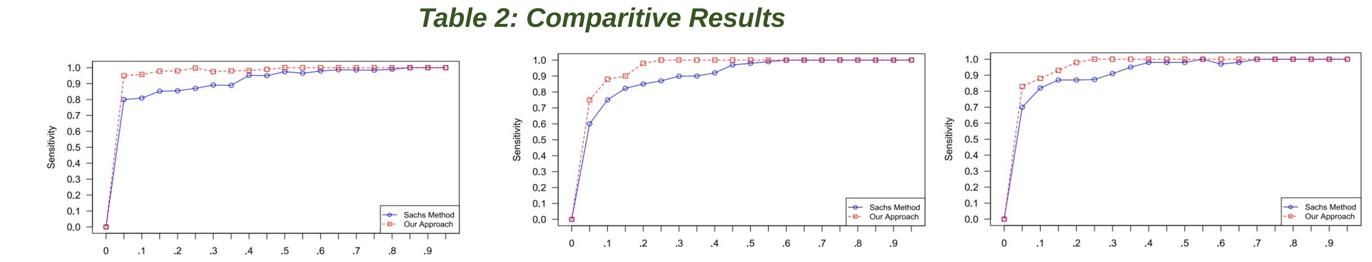
(a) TP: 19, FP: 26, FN: 1 (a) TP: 17, FP: 8, FN: 3 (a) TP: 18, FP: 2, FN: 2

Figure 2: Network Inferred from:(a) PC, (b) GDS, (c) GIES, (d) ICP, (e) simy, (f) Re-implemented Sachs method (g) 'Learn and Vote'

Comparitive Benchmark

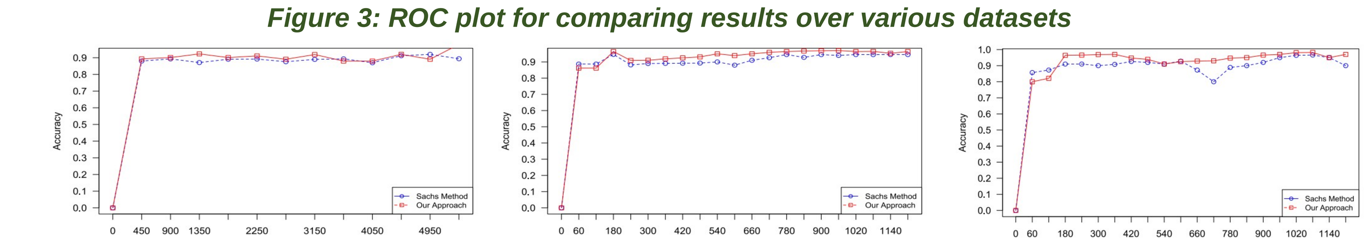
Dataset	Metric	Causal Discovery Algorithms						
		PC	GDS	GIES	ICP	simy	Sachs et al.	Learn and Vote
Flow Cytometry	Precision	0.5714	0.4186	0.377	1	0.4222	0.68	0.89
	Recall	0.4	0.9	0.85	0.45	0.95	0.85	0.89
	F1 Score	0.47	0.572	0.522	0.62	0.584	0.7558	0.89
Lizards	Precision	1	1	1	0	1	1	1
	Recall	1	1	1	0	1	0.5	0.5
	F1 Score	1	1	1	0	1	0.667	0.667
Asia_mut1	Precision	1	0.625	0.625	1	0.31578	0.77	1
	Recall	0.75	0.625	0.625	0.5	0.75	0.875	0.75
	F1 Score	0.857	0.625	0.625	0.666	0.4444	0.8237	0.857
Asia_mut2	Precision	1	0.85714	0.85714	1	0.3043	0.666	1
	Recall	0.75	0.75	0.75	0.5	0.875	0.75	0.75
	F1 Score	0.857	0.8	0.8	0.666	0.4928	0.7058	0.857
gmInt	Precision	0.75	0.889	0.889	1	0.889	0.8571	1
	Recall	0.75	1	1	0.375	1	0.75	0.75
	F1 Score	0.75	0.94	0.94	0.5454	0.94	0.8	0.857
Alarm_mut1	Precision	0.666	0.25	0.26	0.7	n/a	0.625	0.564
	Recall	0.434	0.217	0.26	0.26	n/a	0.4464	0.4
	F1 Score	0.526	0.2325	0.26	0.38	n/a	0.52	0.468
Alarm_mut2	Precision	0.666	0.411	0.5128	0.6	n/a	0.725	0.769
	Recall	0.434	0.456	0.434	0.21	n/a	0.63	0.642
	F1 Score	0.526	0.432	0.47	0.3115	n/a	0.675	0.7
Insurance_mut1	Precision	0.7143	0.36	0.3617	0.7	n/a	0.857	0.8
	Recall	0.288	0.3461	0.327	0.25	n/a	0.577	0.538
	F1 Score	0.4107	0.352	0.3435	0.368	n/a	0.689	0.643
Insurance_mut2	Precision	0.7143	0.355	0.366	0.64	n/a	0.676	0.6857
	Recall	0.288	0.423	0.423	0.21	n/a	0.4423	0.4615
	F1 Score	0.4107	0.386	0.392	0.316	n/a	0.535	0.5517

Table 2: Comparitive Results



(a) Flow cytometry (b) Asia_mut1 (c) Asia_mut2

Figure 3: ROC plot for comparing results over various datasets



(a) Flow cytometry (b) Asia_mut1 (c) Asia_mut2

Figure 4: Sample size vs Accuracy plot for comparing results over various datasets

Conclusion

- We provide a benchmark for Causal learning algorithms in case of mix of observation and experimental dataset.
- We presented a novel way to combine data generated from various conditions (observational or interventional)
- Our results show that while we can combine data from various experiments to learn a Causal Network, such methods can also predict a large number of False Positives.
- Our approach helps reduce the detection of False Positives.

❑ Future Work

- What if targets of intervention are unknown?
- Unequal size of data per experiment
- Categorizing which interventions are more informative
- Presence of latent variable
- Presence of cycles

Oregon State University