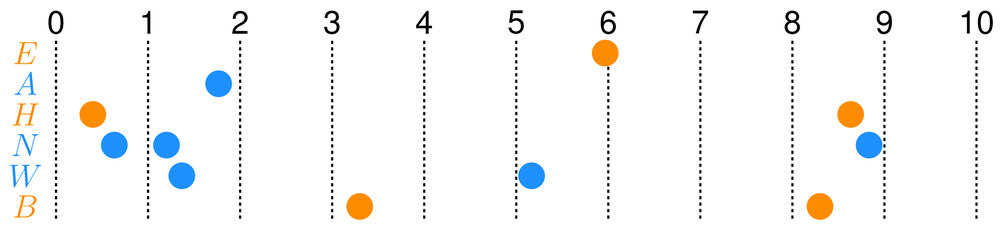
Causal Structure Learning via Temporal Markov Networks

Aubrey Barnard • **David Page**

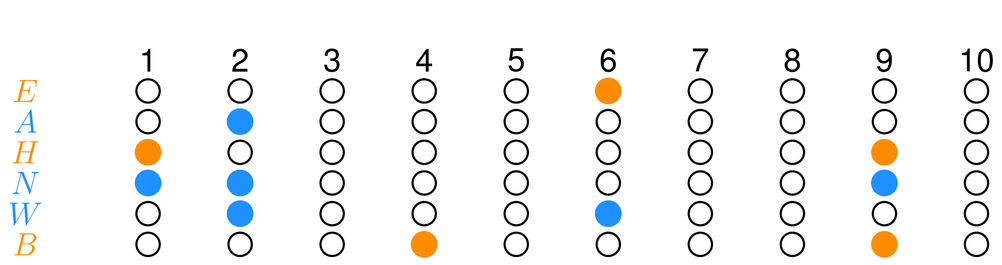
Computer Sciences | Biostatistics and Medical Informatics • University of Wisconsin–Madison

1. Given Timelines

Continuous

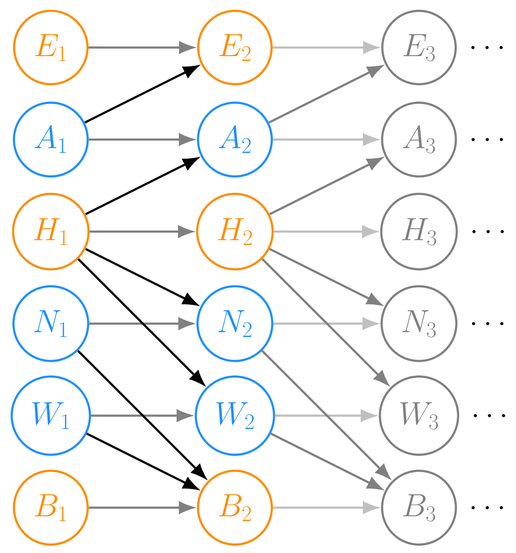


Discretized

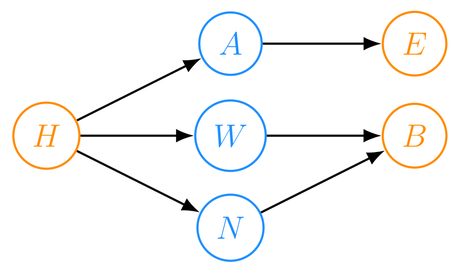


2. Learn Causal Structure

Causal DBN



Summary BN



3. Using Temporal Markov Networks

A temporal Markov network (TMN) is a log-linear model that defines a probability distribution over possible timelines S through hierarchical, temporal features f_i.

$$P(S = s) = \frac{1}{Z} \exp\left(\sum_i \theta_i f_i(s)\right) \quad (1)$$

$$Z = \sum_{s \in S} \exp\left(\sum_i \theta_i f_i(s)\right) \quad (2)$$

The parameters θ_i are estimated with standard maximum likelihood.

Temporal Features

Timelines are modeled with binary feature functions that indicate event (co-)occurrences and temporal relationships.

- *event*, $f_S(x)$: true if event x occurs in S (atemporal)
- *event@*, $f_S(x_t)$: true if event x occurs at t in S (atemporal)
- *co-occur*, $f_S(x, y)$: true if events x and y occur in S (atemporal)
- *co-occur@*, $f_S(x_{t_1}, y_{t_2})$: true if x occurs at t_1 and y occurs at t_2 in S (temporal)
- *before*, $f_S(x \to y)$: true if x and y occur in S and x occurs before y (temporal)
- *before-*δ, $f_S(x_T \to y_{T+\delta})$: true if x and y occur in S and x occurs δ timesteps before y (temporal)
- *before3*, $f_S(\{x, y\} \to z)$: true if x, y, and z occur in S and both x and y occur before z (temporal)

The "@" features are anchored to specific timesteps, but the other features float, tying parameters across timesteps.

Modeling a Stochastic Process

1. Choose features that capture events and the level of interactions.
2. Instantiate the features ($f_S(\cdot) \mapsto f_i(S)$) for each desired combination of events and times.

Depending on the choice of features and the parameter tying they induce, TMNs can represent undirected analogs of BNs, DBNs, HMMs, etc.

Parameters → Independence → Structure

Structure learning via TMNs is enabled by the key idea that $\theta_i = 0$ indicates statistical independence which indicates the absence of an edge in the underlying graph G. Thus, learn the parameters to determine edges and then direct edges with time. With real data, decide zeros by thresholding.

Causal Structure Learning

Structure learning via TMNs will be causal if the distribution of the true stochastic process

- is faithful to a graph G,
- is Markov with respect to G, and
- has no instantaneous interactions.

Advantages

- Non-combinatorial structure learning algorithm
- Smooth, continuous, convex optimization
- Fast, any-time, joint estimation of edges
- Robustness via aggregate sufficient statistics

ADE Discovery Experiments

TMNs were applied to the causal task of discovering adverse drug events (ADEs) in electronic medical records (EMR) data. This task was developed by the Observational Medical Outcomes Partnership (OMOP) as part of an initiative to improve patient health through automatic drug safety surveillance.

Data Sets

OMOP provided real-world EMR (GE) and claims (CCAE, MDCD, MDCR, MSLR) databases.

Methods

PC Standard causal BN structure learning via independence testing

TMN-PC TMN with static ("@") BN features

BNF-DBN DBN learning with BNFinder (optimal search and score)

TMN-DBN TMN with first-order DBN features

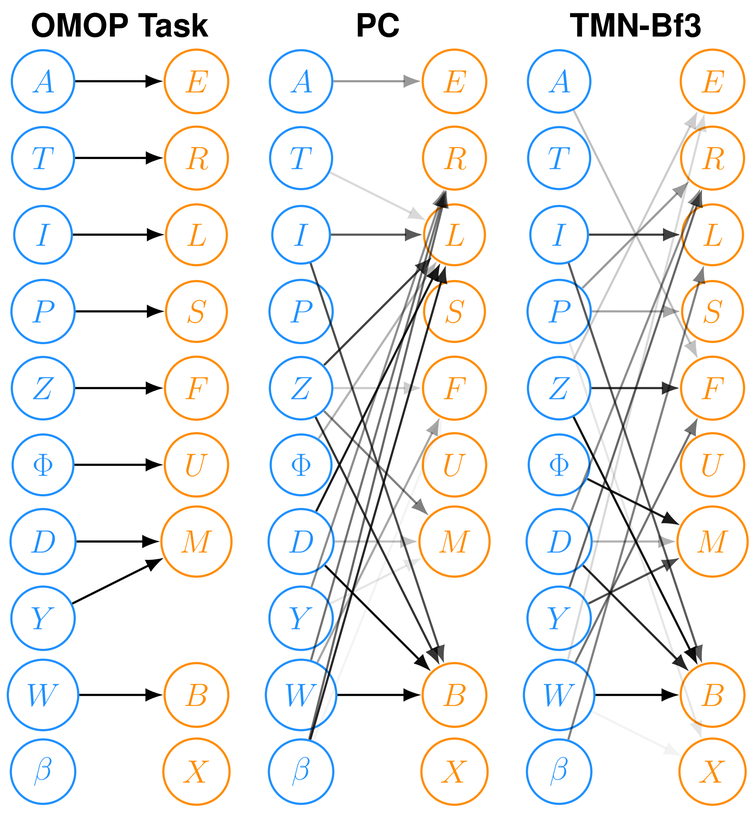
TMN-Bf3 TMN with DBN features plus long-range and 3-way *before* features

Random guessing

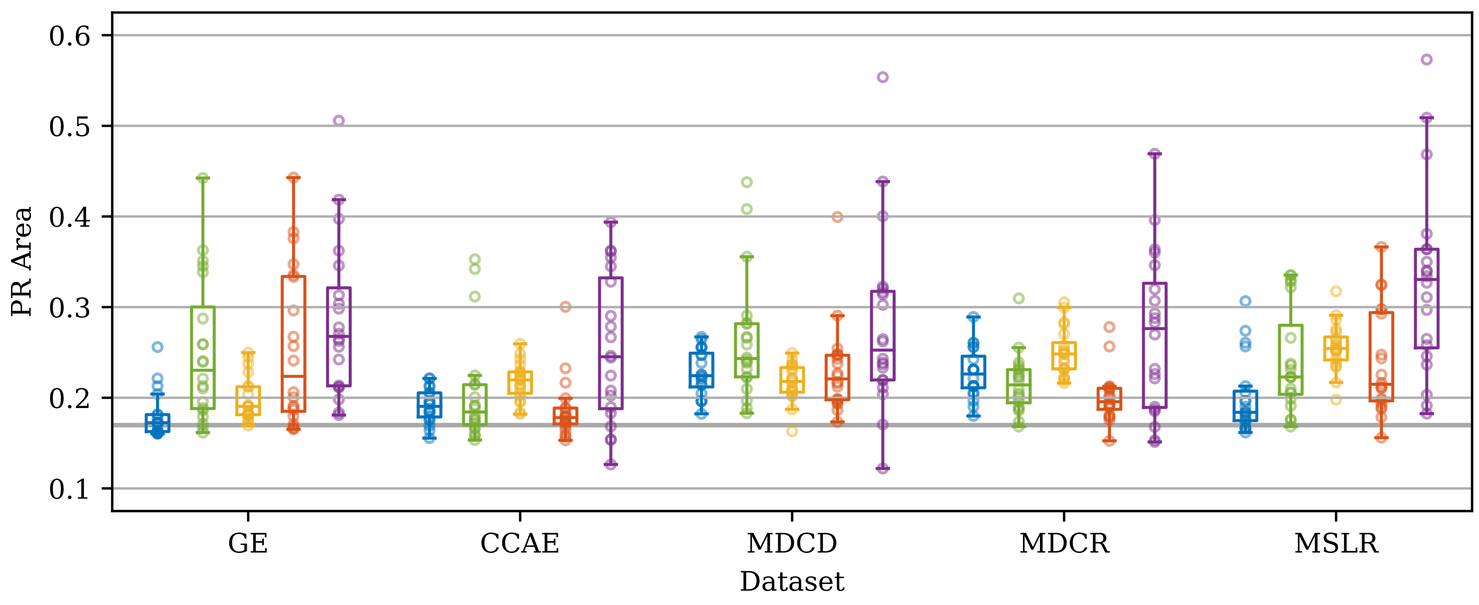
OMOP Task & Events of Interest

Drug	Condition	ADE?
A ACE inhibitors	E angioedema	+
T amphotericin B	R acute renal failure	+
I antibiotics	L acute liver failure	+
P antiepileptics	S aplastic anemia	+
Z benzodiazepines	F hip fracture	+
Φ bisphosphonates	U upper GI ulcer	+
D tricyc. antidep.	M acute MI	+
Y typ. antipsycho.	M acute MI	+
W warfarin	B bleeding	+
β beta blockers	X MI mortality	−
N NSAIDs	H hypertension	

Results: Learned Graphs



Results: ADE Recovery



Thanks to the Center for Predictive Phenotyping (NIH BD2K Initiative grant U54 AI117924, NIGMS grant 2RO1 GM097618) and IMEDS for supporting this work.