Typing Assumptions Improve Identification in Causal Discovery


Philippe Brouillard [1,2]


Perouz Taslakian [1]


Alexandre Lacoste [1]


Sébastien Lachapelle [2]


Alexandre Drouin [1]

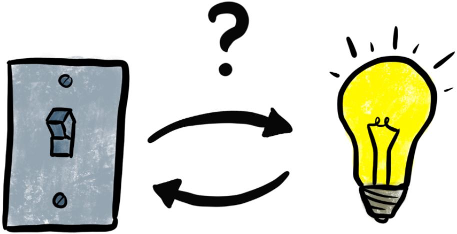
[1] Element AI, a ServiceNow company
[2] Mila & DIRO, Université de Montréal

Mila Université de Montréal

ELEMENT AI
a ServiceNow Company

Motivation

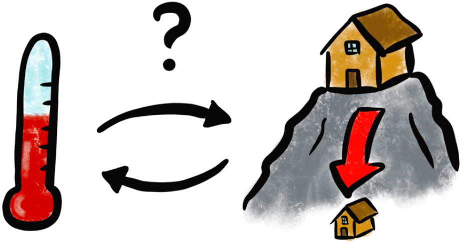
Can a light bulb change the state of a switch?



Mila Université de Montréal
ELEMENT AI
a ServiceNow Company

Motivation

Can the temperature of a city alter its altitude?



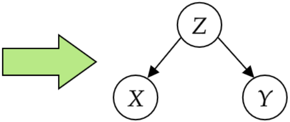
Motivation

- As humans, we understand the implausibility of causal relationships between certain types of entities.
- In fact, most of the time, we use prior knowledge to generalize causal relation between similar entities.
  [Griffiths et al., 2011, Schulz & Gopnik, 2004, Gopnik & Sobel, 2000]

Causal Discovery

The task of causal discovery consists of learning the structure of G based on observations from P_X.

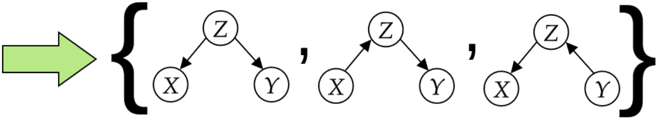
X	Y	Z
1.21	1.58	0.33
1.50	1.84	0.51
$\vdots$	$\vdots$	$\vdots$
0.96	1.07	0.11



Markov Equivalence Class

Unfortunately, from observational data, one can only retrieve a set of equivalent DAGs called the **Markov Equivalence Class** (MEC).

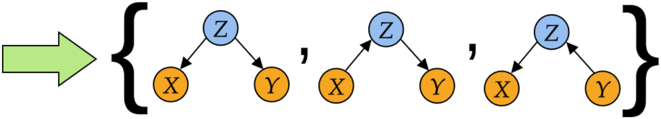
X	Y	Z
1.21	1.58	0.33
1.50	1.84	0.51
$\vdots$	$\vdots$	$\vdots$
0.96	1.07	0.11



Variable types

In our new setting, each vertex has an associated type (that might have been given by an expert). We call these graphs **t-DAGs**.

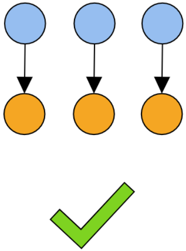
X	Y	Z
1.21	1.58	0.33
1.50	1.84	0.51
⋮	⋮	⋮
0.96	1.07	0.11

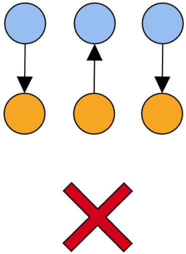


Type consistency: constraints on type interactions

Assumption: all edges between a pair of types are oriented in the same direction.

We call t-DAGs that satisfy this condition **consistent t-DAGs**.

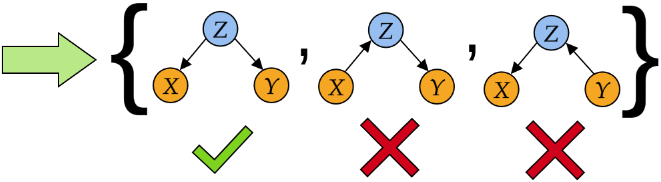




t-MEC: an equivalence class for consistent t-DAGs

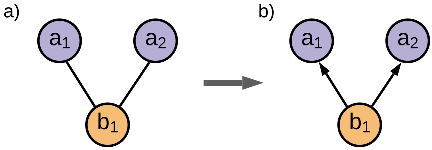
With this assumption, the size of the MEC can be greatly reduced by removing t-DAGs that violate type consistency. We call this equivalence class a **t-MEC**.

X	Y	Z
1.21	1.58	0.33
1.50	1.84	0.51
⋮	⋮	⋮
0.96	1.07	0.11



Theorem: identification guarantees for random graphs

- We show that there exists conditions under which our assumptions lead to benefits in identification.
- For a random sequence of t-DAGs with a fixed number of types, the size of the t-MEC **converge to a singleton** exponentially fast with **the number of vertices** (measured variables).
- **Proof sketch:** There exists a structure, called a two-type fork, that forces the orientation of edges due to type consistency. As the graph grows, the probability of observing it converges to 1 and thus all edges of the t-DAG are oriented.

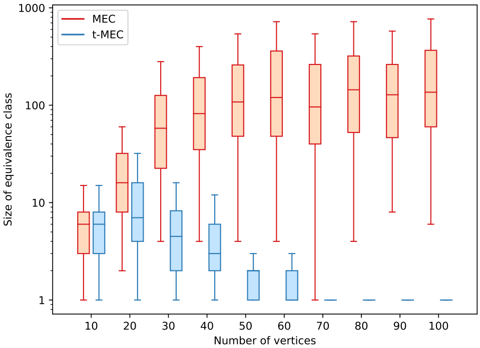


Two-type fork

Mila Université de Montréal
ELEMENT AI
a ServiceNow Company

Experiments

As the number of vertices increase, |t-MEC| decreases while |MEC| increases.



Conclusion

This work shows that our typing assumptions can help reduce the size of the MEC and thus help in causal discovery when our assumptions hold.

Future work:

- Explore practical applications (tiered background knowledge [Andrews, 2020])
- Relax our assumption for more realistic settings
- Learn types automatically from metadata related to each variable

Mila Université de Montréal

ELEMENT AI
a ServiceNow Company

Thank you!





Philippe Brouillard [1,2]

Perouz Taslakian [1]

Alexandre Lacoste [1]

Sébastien Lachapelle [2]

Alexandre Drouin [1]

Come to our poster to learn more about:

- our assumptions based on variable types.
- a simple algorithm to incorporate these assumptions in causal discovery.
- the theoretical result that guarantee convergence of the equivalence class to a singleton.

Mila Université de Montréal
ELEMENT AI
a ServiceNow Company

References

Andrews, B. (2020).
On the completeness of causal discovery in the presence of latent confounding with tiered background knowledge.
In S. Chiappa & R. Calandra (Eds.), *The 23rd International Conference on Artificial Intelligence and Statistics, AISTATS 2020, 26-28 August 2020, Online [Palermo, Sicily, Italy]*, volume 108 of *Proceedings of Machine Learning Research* (pp. 4002–4011).: PMLR.

Gopnik, A. & Sobel, D. M. (2000).
Detecting blickets: How young children use information about novel causal powers in categorization and induction.
Child development, 71(5), 1205–1222.

Griffiths, T. L., Sobel, D. M., Tenenbaum, J. B., & Gopnik, A. (2011).
Bayes and blickets: Effects of knowledge on causal induction in children and adults.
Cognitive Science, 35(8), 1407–1455.

Schulz, L. E. & Gopnik, A. (2004).
Causal learning across domains.
Developmental psychology, 40(2), 162.

Mila Université de Montréal

ELEMENT AI
a ServiceNow Company