Structural Causal Modelling

Outline

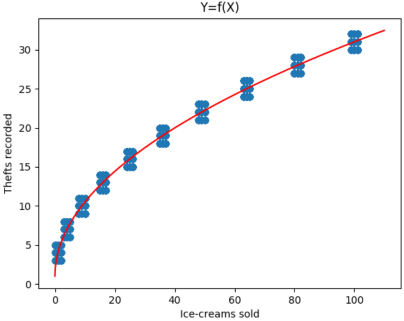
1. Review of Causality
2. Structural Causal Modelling
3. Computing Causal Quantities
4. Conclusions

2. Review of Causality

Review: Ice Creams and Thefts

Assume we monitored the number of *ice-creams sold* (Ice) and the number of *thefts* (Thf) in our town:

Ice	Thf
36	20
35	18
101	31
17	12
50	23
...	...

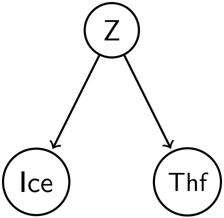


- From data we can learn to *predict*, but not to *control*.
- *Correlation is not causation*.

Review: Common Causes

Assume causality is *directed* and *mechanistic*.

We explain correlation through a *common cause* (Z), such as the season or temperature:



- Instance of the more general *Reichenbach principle*.

Review: Causal Questions

We want to ask questions *beyond standard statistics/ML*.

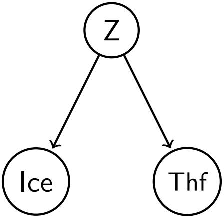
Causal questions exist at different levels:

L3. Counterfactuals	What would have Thf been, had Ice been set to 0 when instead it was observed to be 42?
L2. Causal Effects	What is the effect on Thf of forcing Ice to 0?
L1. Observational Relationships	What is the probability of Thf observing Ice=0?

- This defines *Pearl's causality ladder*.

3. Structural Causal Modelling

Why Graphical Models?



Provides a clear and visual *expression of assumptions.*

They intuitively fit two assumptions:

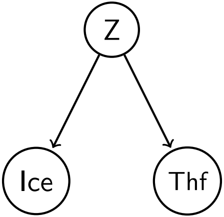
(i) *Causal directionality*

(ii) *Mechanism*

From BNs to Causal Models

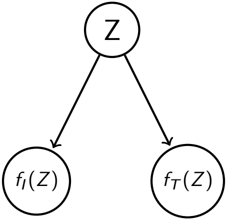
Bayesian network models

$P(Ice, Thf, Z) =$
$\mathbf{P}(\mathbf{Thf}|\mathbf{Z})\mathbf{P}(\mathbf{Ice}|\mathbf{Z})\mathbf{P}(\mathbf{Z})$



Mechanistic causal modelling

$P(Ice, Thf, Z) =$
$\mathbf{P}(\mathbf{Thf}|\mathbf{Z})\mathbf{P}(\mathbf{Ice}|\mathbf{Z})\mathbf{P}(\mathbf{Z})$



We extend *Bayesian networks* to express (i) *causal directionality* and (ii) *mechanisms*.

Structural Causal Models

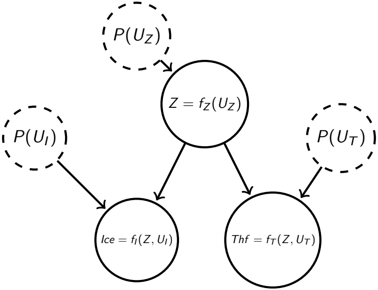
A **structural causal model** (SCM) is defined as a tuple:

$$\mathcal{M} = \langle \mathcal{X}, \mathcal{U}, \mathcal{F}, \mathcal{P} \rangle$$

where:

- $\mathcal{X}$ is a set of *endogenous nodes* (*variables of interest*);
- $\mathcal{U}$ is a set of *exogenous nodes* (*noise*);
- $\mathcal{F}$ is a set of *structural functions*, one for each endogenous node;
- $\mathcal{P}$ is a set of *probability distributions*, one for each exogenous node.

SCM: Remarks



Remarks:

- SCM implies a DAG
- We assume *acyclicity* and *independent noise*

What can we use a SCM for?

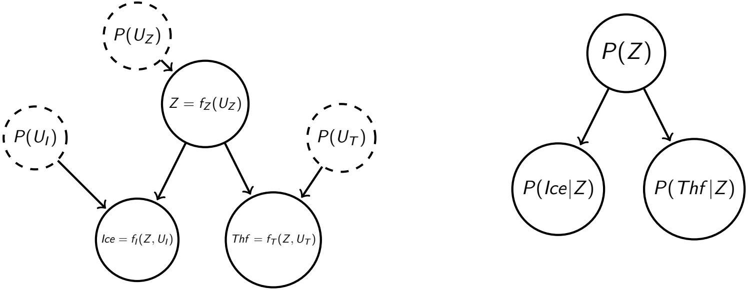
Given a *completely specified* SCM $\mathcal{M}$ to answer questions on any rung of **Pearl's ladder**.

L3. Counterfactuals	What would have Thf been, had Ice been set to 0 when instead it was observed to be 42?
L2. Causal Effects	What is the effect on Thf of forcing Ice to 0?
L1. Observational Relationships	What is the probability of Thf observing Ice=0?

4. Computing Causal Quantities

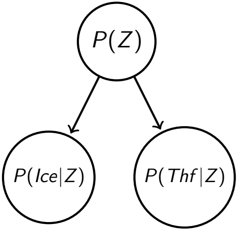
L1: SCM and Observational Questions

A *pushforward* of the distributions on the exogenous nodes defines a distribution over all the endogenous variables.



We reduce the SCM to a Bayesian network and compute $P_{\mathcal{M}}(Ice, Thf, Z)$ or any related quantity.

L1: Example



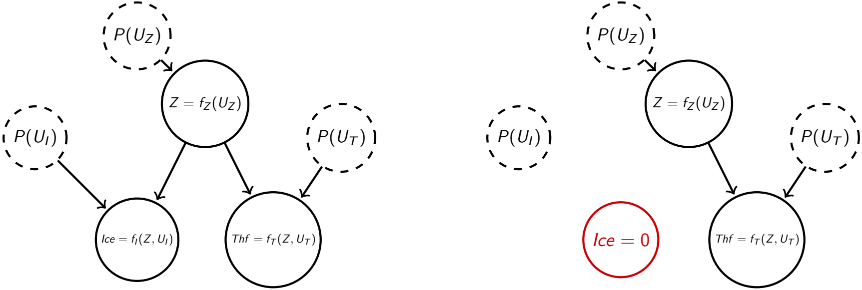
What is the probability of Thf observing Ice=0?

$$P_{\mathcal{M}}(Thf|Ice=0)=$$
$$=\frac{P_{\mathcal{M}}(Thf,Ice=0)}{P_{\mathcal{M}}(Ice=0)}=$$
$$=\frac{\sum_{Z}P_{\mathcal{M}}(Thf,Ice=0,Z)}{\sum_{Thf,Z}P_{\mathcal{M}}(Thf,Ice=0,Z)}$$

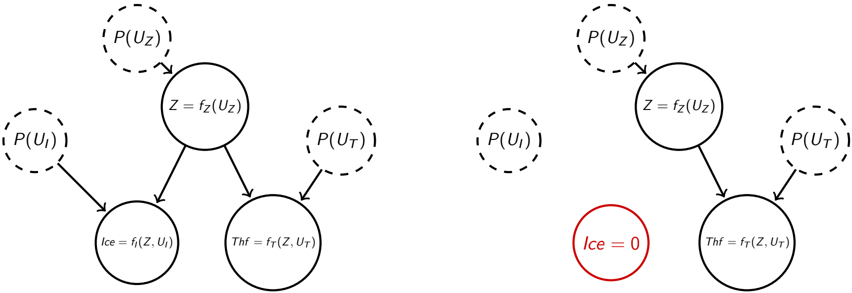
L2: Interventions

We define an **intervention** operator $do(X = x)$ which replaces the function in variable X with the constant x.

$$do(Ice = 0)$$



L2: Interventions - Remarks



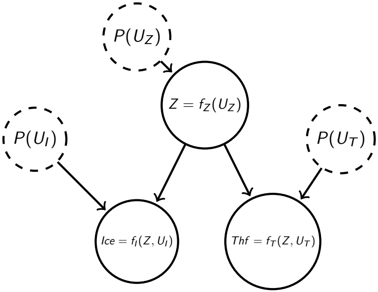
Remarks:

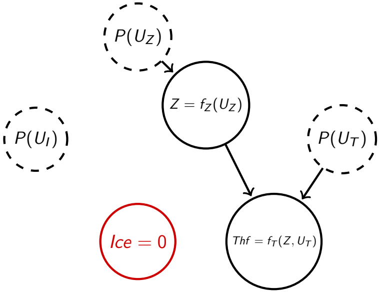
- We assume interventions to be *exact* and *local*.
- An intervention $do(X = x)$ on model $\mathcal{M}$ defines a new *post-intervention model* $\mathcal{M}_{do}$.
- From a SCM we can generate multiple post-interventional models.

L2: SCM and Interventional Questions

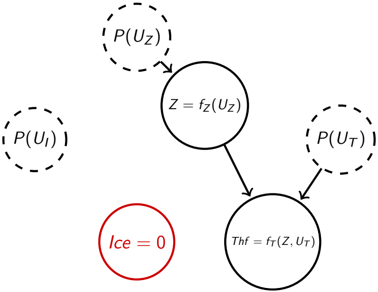
We answer an interventional question in the new model defined by an intervention:

$$P_{\mathcal{M}}(Thf|do(Ice = 0)) = P_{\mathcal{M}_{do}}(Thf|Ice = 0)$$





L2: Example

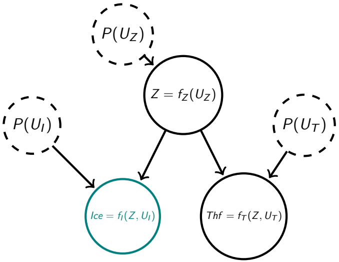


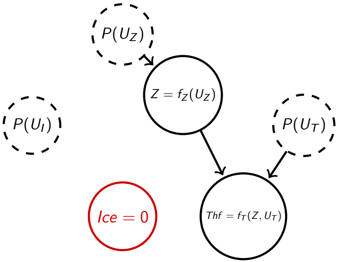
What is the effect on Thf of forcing Ice to 0?

$$P_{\mathcal{M}}(Thf|do(Ice = 0)) = P_{\mathcal{M}_{do}}(Thf|Ice = 0) = P_{\mathcal{M}_{do}}(Thf)$$

L2: Conditioning and Intervening

Conditioning $\neq$ Intervention





$$P_{\mathcal{M}}\left(Thf \,|\, Ice = 0\right)$$

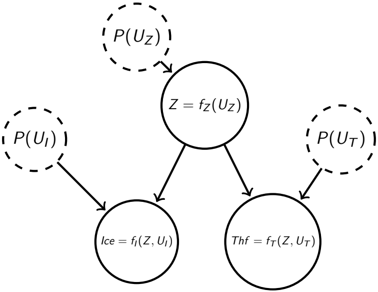
$$P_{\mathcal{M}}\left(Thf \,|\, do\left(Ice = 0\right)\right)$$
$$P_{\mathcal{M}_{do}}\left(Thf \,|\, Ice = 0\right)$$

Knowledge of $Ice = 0$ allows inference on distribution of Z and then Thf.

Knowledge of $do\left(Ice = 0\right)$ does not affect the distribution of Z.

L3: Counterfactuals

A **counterfactual** is a possible outcome which we did not observe, but could have happened under different circumstances.

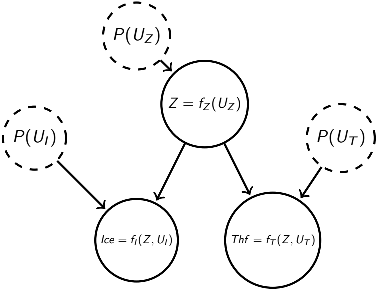


Given that we observed *Thf* when *Ice* happened to be 42, what would have *Thf* been if we had set *Ice* to 0?

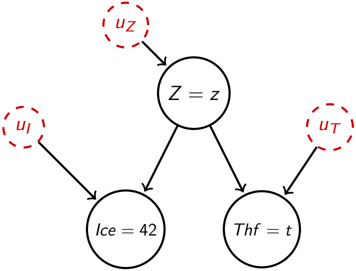
$$\cancel{P_{\mathcal{M}}(Thf \mid Ice = 42, do(Ice = 0))} \qquad P_{\mathcal{M}}(Thf_{Ice=42} \mid do(Ice = 0))$$

L3: Computing Counterfactuals

1. *Abduction:* we start from the factual world, and infer the value/distribution of exogenous variables from observations.



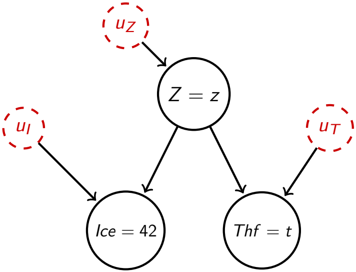
(Original model)



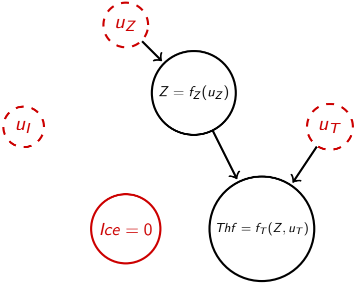
(Abduction)

L3: Computing Counterfactuals

2. *Action:* intervene as requested in the counterfactual.



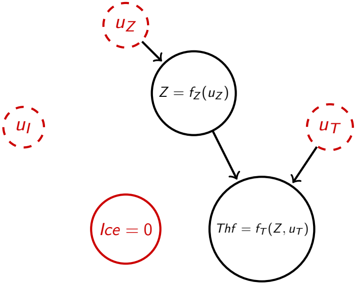
(Abducted model)



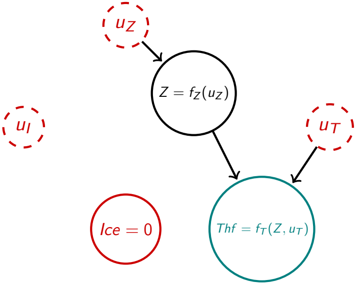
(Action)

L3: Computing Counterfactuals

3. *Prediction:* compute the variable of interest in the counterfactual model.

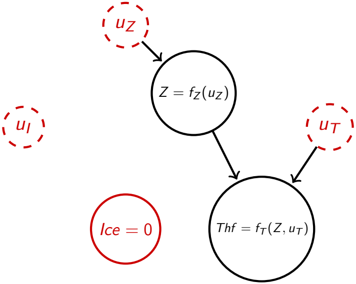


(Abducted-Acted model)



(Prediction)

L3: Example

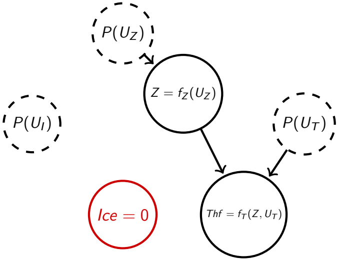


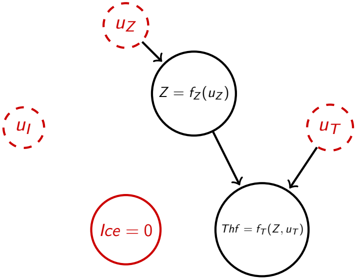
What would have Thf been, had Ice been set to 0 when instead it was observed to be 42?

$$P_{\mathcal{M}}(Thf_{Ice=42} | do(Ice = 0))$$
$$P_{\mathcal{M}_{Ice=42_{do}}}(Thf | Ice) =$$
$$P_{\mathcal{M}_{Ice=42_{do}}}(Thf)$$

L3: Intervening and Counterfactuals

Intervention $\neq$ Counterfactuals





$$P_{\mathcal{M}}\left(Thf | do\left(Ice = 0\right)\right)$$
$$P_{\mathcal{M}_{do}}\left(Thf | Ice = 0\right)$$

The value of *Thf* is independent of *Ice*, determined by a random samples of *Z*.

$$P_{\mathcal{M}}(Thf_{Ice=42} | do(Ice = 0))$$
$$P_{\mathcal{M}_{Ice=42do}}(Thf | Ice = 0)$$

The value of *Thf* is always independent of *Ice*, but the sample from *Z* is conditioned to the value that produced $Ice = 42$.

5. Conclusion

Recap

1. *Causal queries* require a formalism beyond (observational) statistics.
2. SCMs provide a new *rigorous formalism* for causal queries.
    - Graphical formalism
    - BN $\rightarrow$ SCM
3. SCMs give rigorous form and answers to *causal queries*.
    - *Observational questions:* reduction of SCM to BN
    - *Causal effects:* intervention and reduction to post-intervention model
    - *Counterfactuals:* counterfactual algorithm and reduction to counterfactual model

Further directions

"Given a *completely specified* SCM $\mathcal{M}$...", what if not?

- Knowledge of DAG plus data $\rightarrow$ **Causal inference**
  - How to derive causal quantities from observational data?
- Only data $\rightarrow$ **Causal discovery**
  - How to learn a DAG from data?

Other assumptions may be dropped or added:

- Dropping assumption on independent noise $\rightarrow$ *Semi-Markovian SCMs*
- Dropping assumption on exactness of interventions $\rightarrow$ *Soft interventions*
- Adding assumptions on the form of functions $\rightarrow$ *Parametric causality*

Further directions

Causality has deep intersections with AI and ML:

- *Causal and Anti-Causal Learning*: exploiting causal directions in learning.
- *Explainability and Fairness*: relying on causal relationship to explain and justify results.
- *Causal Reinforcement Learning*: endowing agents with causal reasoning.
- *Causal Representation Learning*: generating representations with causal meaning.

Questions?

Any questions?

(Feel free to get in touch at any time at `fabio.zennaro@warwick.ac.uk` or `fm.zennaro@gmail.com`)

Resources

- *Reproduce and play with the model discussed*:
  - *Notebooks*, e.g.: `https://github.com/FMZennaro/CausalInference`
  - *Libraries*, e.g.: `https://pgmpy.org/`, `https://github.com/py-why/dowhy`

- *Foundational books on SCMs*:
  - J. Pearl, and D. Mackenzie. *The Book of Why*
  - M. Glymour, J. Pearl, and N.P. Jewell. *Causal inference in statistics: A primer*
  - J. Pearl. *Causality*
  - J. Peters, D. Janzing, and B. Schölkopf. *Elements of causal inference: foundations and learning algorithms*

Further resources:

- *Causal discovery:* C. Glymour, K. Zhang, P. Spirtes, *Review of causal discovery methods based on graphical models*
- *Causal learning:* B. Schölkopf, D. Janzing, J. Peters, E. Sgouritsa, K. Zhang, J. Mooij. *On causal and anticausal learning*.
- *Causal fairness:* M.J. Kusner, J. Loftus, C. Russell, R. Silva. *Counterfactual fairness*
- *Causal reinforcement learning:* `https://crl.causalai.net/`
- *Causal representation learning:* B. Schölkopf, F. Locatello, S. Bauer, N.R. Ke, N. Kalchbrenner, A. Goyal, Y. Bengio. *Toward causal representation learning*
- *Causal abstractions:* `https://github.com/FMZennaro/CausalAbstraction`

From distributions to BNs to CBNs

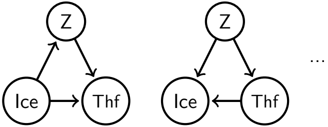
Statistical modelling

$\mathbf{P}(\mathbf{Ice}, \mathbf{Thf}, \mathbf{Z}) =$

$P(Thf|Z, Ice)P(Z|Ice)P(Ice)$

$P(Ice|Thf, Z)P(Thf|Z)P(Z)$

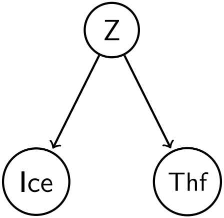
...



Bayesian network modelling

$P(Ice, Thf, Z) =$

$\mathbf{P}(\mathbf{Thf}|\mathbf{Z})\mathbf{P}(\mathbf{Ice}|\mathbf{Z})\mathbf{P}(\mathbf{Z})$

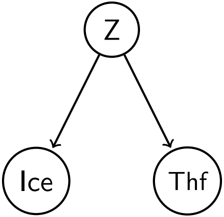


Causal Bayesian networks allows to express causal *directionality* (i).

From CBNs to SCMs

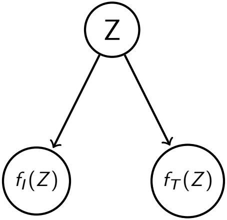
Causal Bayesian network modelling

$P(Ice, Thf, Z) =$
$\mathbf{P(Thf|Z)P(Ice|Z)P(Z)}$



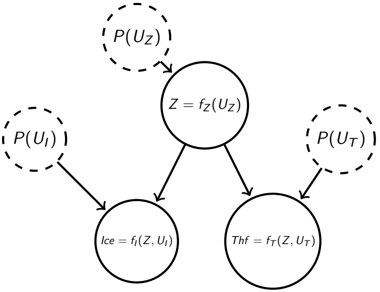
Mechanistic causal modelling

$P(Ice, Thf, Z) =$
$\mathbf{P(Thf|Z)P(Ice|Z)P(Z)}$



We need to extend *causal Bayesian networks* to express *mechanisms* (ii).

Assumptions about SCMs



1. **Acyclicity**: causality is acyclic.
2. **Causal arrows**: an arrow is a causal link, a missing arrow means no causal link.
3. **Causal relationship completeness**: all causes among the variables in the model are present.
4. **Common cause completeness / Independent noise**: all common causes are modeled.

Causal Markov assumption: a node is independent of its non-effects given its direct causes (consequence of 1 and 4).