Introduction to causal models. Lecture 1

Tsinghua Logic Summer School

Tadeg Quillien

University of Edinburgh

2026

Introduction

Housekeeping

(Slides are available online)

Instructors:

- Dr Tadeg Quillien (tadeg.quillien@ed.ac.uk)
- Dr Bonan Zhao (b.zhao@ed.ac.uk)

Teaching assistants:

- Xiangyuan Ji (xy-ji23@mails.tsinghua.edu.cn)
- Wenlong Zheng (zhengwl21@mails.tsinghua.edu.cn)

Time:

- 9:50-12:15 Monday to Friday.
- We take a 10-minute break at $\sim$10:40 and 11:30.

Assessment

Three homework assignments:

- Released on Tuesday, Wednesday, Friday.

One final exam:

- In-class, closed-book pen-and-paper exam
- 45 min, on Friday ($\sim$11:30).

AI policy:

- AI use to help with assignments is NOT permitted.

Is smoking harmful for your health?

- Actual question asked by 20th century researchers!

Problem:

- Data showed that smokers have more lung cancer,
- But maybe this had a benign explanation,
    - E.g. there is a gene that causes smoking AND lung cancer.

Illustrates a major unresolved intellectual problem at the time:

- We are interested in whether something **causes** something else.
- But the notion of causality appears nowhere in the formalism of statistics (or any other formalism)!

Early ideas for formalizing causality

Two options seemed promising:

- Maybe statements about causality can be reduced to statements about statistics and probability (e.g. Reichenbach, 1956).
- Maybe causality is an archaic notion that has no place in our scientific theorizing (e.g. Russell, 1912).

A surprising development in late 80s / early 90s: *neither option is correct!*

- We can have a respectable formal understanding of causality,
- But it is provably not reducible to pure statistics and probability.

This surprising discovery is now known as the **causal revolution**.

A brief history of causality

- Long-standing interest among philosophers (Aristotle, Hume, Kant...).
- Early/mid-20th century: development of inferential statistics and randomized experiments (Fisher, Neyman, Pearsons...).
  - But no formal notion of causality.
- Early forays into causal modeling in mid-20th century in genetics and economics (Wright, Haavelmo...).
- In late 1980s-early 1990s: Development of the formal framework of causal models (Pearl, Spirtes, Glymour...).
  - Key notion: **intervention**.

Today: the uses of causal models

- Computer science: causal discovery, machine learning (explainability, etc),
- Philosophy: philosophy of science, metaphysics (Jim Woodward, L.A. Paul...),
- Statistics, empirical sciences: estimation of causal effects,
- Linguistics: causality in language,
- Psychology: causal reasoning (Gopnik, Griffiths, Cheng, Waldmann, Lagnado...)

Goal of the course

- Introduction to the foundational ideas driving the causal revolution,
- with some pointers to applications.

Course difficulty / level of formalism:

- Math is overall simple,
- Emphasis is not on mathematical rigor for its own sake,
- **BUT: content builds upon itself.**
  - Easy to get lost if you don't grasp material from an earlier part of the course.

Course outline

- Today: introduction to causal models
- Tuesday: causal reasoning under uncertainty
- Wednesday: program induction
- Thursday: approximate inference
- Friday: counterfactual reasoning, conclusion

Structural causal models

A simple model

- We want to build a very minimal mathematical model of this system (abstracting away from physical details).
- Information to include: When the Switch is on, it turns the Light on.
- Two boolean variables:
  - $S = 1$: the Switch is on
  - $L = 1$: the Lightbulb is on
- An equation: $L = S$
  - (the Lightbulb is on iff the Switch is on.)

A simple model

- An equation: $L = S$
- Problem: the equation can be re-arranged as $S = L$
- But intuitively, there is an asymmetry: the Switch causes the Lightbulb to be on.
    - Not vice-versa.
- Let us mark the asymmetry by replacing the equality sign: $L := S$

Definition

- A **Structural Equation** is an equation with the ':=' sign.
- for any two X, Y, $X := Y$ is NOT equivalent to $Y := X$.

This is a bit like variable declaration in a computer program.

- $a = 2$ cannot be replaced with $2 = a$.

Interventions

Intuitively, the difference between the Switch and the Lightbulb is as follows:

- If we turn off the Switch, the Lightbulb turns off.
- If we turn off the Lightbulb (e.g. by breaking it), this does NOT affect the Switch.

But so far, we don't have a formal way to represent external changes imposed on the system (e.g. breaking the Lightbulb).

- This is where the formal notion of **Intervention** comes in.

Interventions

Definition

An **Intervention** consists in changing a structural equation by replacing the Right-Hand-Side with a new value.

- e.g. we represent 'breaking the lightbulb' by the following intervention:
  - old structural equation: $L := S$
  - new structural equation: $L := 0$

The Do operator

We often represent interventions using the Do() operator. E.g. Do($L = 0$) means that we intervene to set the structural equation for L to be $L := 0$.

Does this give us what we want?

We begin with the system in the following state:

- $S := 1$, $L := S = 1$

Does turning off the Switch turn off the Lightbulb?

- $\text{Do}(S = 0) \to S := 0$
- $\text{L} := \text{S} = 0$ ✓

Does breaking the Lightbulb leave the Switch unchanged?

- $\text{Do}(L = 0) \to L := 0$
- $S := 1$ ✓

That's it!

We have successfully formalized the asymmetry of causality!

Experiments as a way to exploit interventions

In empirical science the gold standard to discover causal relations is to perform experiments. E.g., you want to know if a new drug cures headaches:

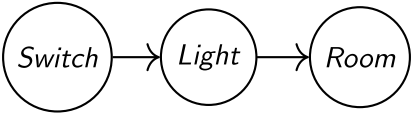
- Randomly assign 100 people to drug
- Randomly assign 100 people to placebo
- Do people in the first group have fewer headaches?

The random assignment consists in an intervention:

- For each participant you perform either $\mathrm{Do}(D = 1)$ or $\mathrm{Do}(D = 0)$.
- If people in the first group have much fewer headaches, you know that the drug has a causal effect.

Structural causal models

Suppose we have several causal relations. E.g. the Lightbulb being on causes the room to be lit:

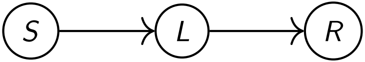


We represent this system with two structural equations:

- $L := S$
- $R := L$

We call such a system of structural equations a **Structural Causal Model** (SCM).

Using Structural Causal Models



- $L := S$
- $R := L$

Definition

If a variable X is in the structural equation of a variable Y, we say that X is a **direct parent** of Y.

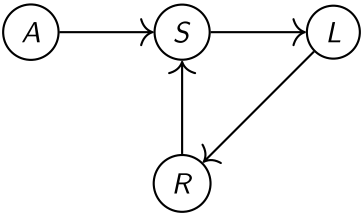
Procedure for determining the state of variables in an SCM

- Determine the state of the variables with no direct parent (more on this later),
- Repeat the following until all variables have a value:
  - Compute the value of variables for whom you know the state of all direct parents.

No cycles!

Procedure for determining the state of variables in an SCM

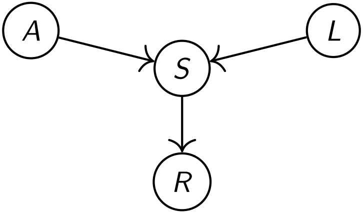
- Determine the state of the variables with no direct parent (more on this later),
- Repeat the following until all variables have a value:
  - Compute the value of variables for whom you know the state of all direct parents.



Some terminology

Definition

- A variable with no direct parent is **exogenous**.
- All other variables are called **endogenous**.



Structural Causal Models

Definition

A Structural Causal Model (SCM) is a tuple $< \mathcal{U}, \mathcal{V}, \mathcal{F}, P(\mathcal{U}) >$:

- A set $\mathcal{U}$ of exogenous variables,
- A set $\mathcal{V}$ of endogenous variables,
- A set $\mathcal{F}$ of structural equations,
- A probability distribution $P(\mathcal{U})$ over exogenous variables.

(The distribution $P(\mathcal{U})$ is how we determine the value of the exogenous variables—more on this later!)

Models and Graphs

We often use graphs to illustrate a causal model:

$$A \rightarrow B \leftarrow C$$

These graphs are called **Directed Acyclic Graphs** (DAGs).

Remark

The same DAG can be used to illustrate different SCMs. For example:

- $B := A \vee C$
- $B := A \wedge C$

So a DAG is an incomplete representation of an SCM.

The ladder of causation

There are different kinds of questions we can ask on the basis of an SCM:

- If we see that X=x, then what is the value of Y?
  - $\rightarrow$ Level 1: **observational** query
- If we $\text{Do}(X = x)$, what is the value of Y?
  - $\rightarrow$ Level 2: **interventional** query
- If X had been x (instead of the value it currently has), what would have been the value of Y?
  - $\rightarrow$ Level 3: **counterfactual** query

The three question classes form a 'hierarchy'.

- SCMs can be used to answer all three kinds of questions,
- Some other models can only answer Level-1 and Level-2 questions,
- Yet some other models can only answer Level-1 questions.

Example: observational vs interventional queries

$L := S$

- If we see $L = 0$, we learn that $S = 0$.
- If we $\mathrm{Do}(L = 0)$, this doesn't tell us anything about S.

If we used normal equations instead of structural equations, we wouldn't be able to tell the difference (we'd be stuck at Level 1 of the ladder).

What about counterfactual queries?

See lecture on Day 5.

The story so far

- We can formalize the asymmetry of causality using **Interventions**.
- Equations with respect to which interventions are well-defined are called **Structural Equations**.
- A causal system can be modeled as a system of Structural Equations, or **Structural Causal Model (SCM)**.

Outstanding questions

- So far all causal systems have been deterministic
- How can we represent uncertainty and stochasticity?