Interrogating Your Twin

Causal Reasoning in Manufacturing Systems

Prasanna Bhogale
prasanna@romulan.ltd

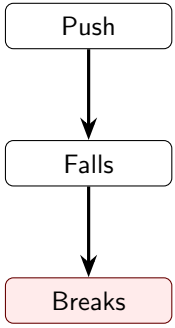
The Fifth Elephant 2026 Pune Edition

ROMULAN

Every three-year-old is a causal reasoner

Children reason causally before they can read:

- "The **wind** knocked it over" (not: wind and falling co-occurred)
- "If I **push** this, it will fall" (intervention)
- "It **would not** have broken if you hadn't dropped it" (counterfactual)



Observation $\neq$ Intervention

What statistics gives you:

$$P(Y \mid X)$$

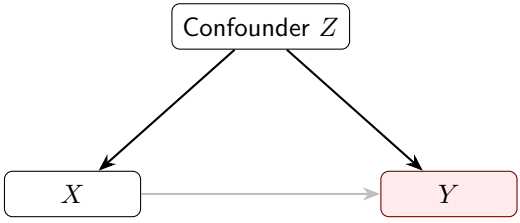
"What is the probability of Y, given that we *observed* X?"

What you actually need:

$$P(Y \mid \mathbf{do}(X))$$

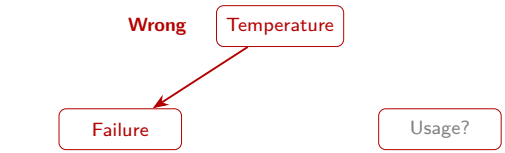
"What is the probability of Y, if we *force* X?"

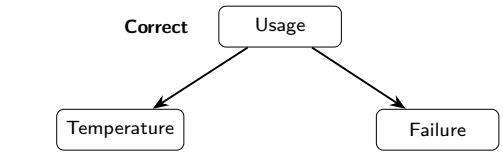
Statistics lacks the notation to express this difference.



X and Y both rise and fall together...
but only because Z moves them both.

The tool: Directed Acyclic Graphs (DAGs)





A **DAG** is a drawing of your causal beliefs:

Node = a variable you measure

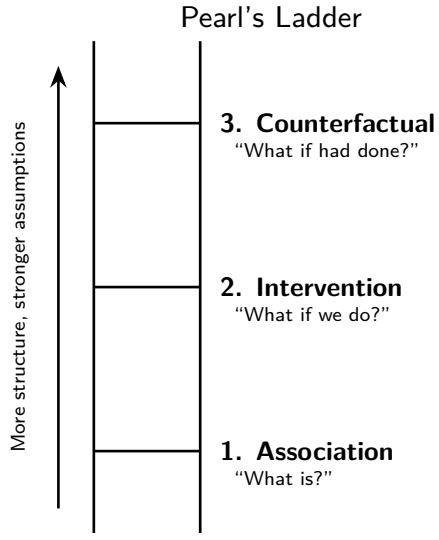
Arrow = "A causes B"

No arrow = "A does *not* cause B"

Pearl & Mackenzie, *The Book of Why* (2018)

What is in this talk

1. Why predictive maintenance models can **mislead** you
2. Pearl's Ladder of Causation — a **framework**
3. A causal inference workflow with **real factory data**
4. Where this fits in your Industrie 4.0 stack



Running example: a manufacturing facility

Machine-shifts	$\sim 1{,}500$
Distinct machines	~ 23
Shifts per day	2 (12-hour)
Observation window	47 days
Breakdown rate	$\sim 3\%$

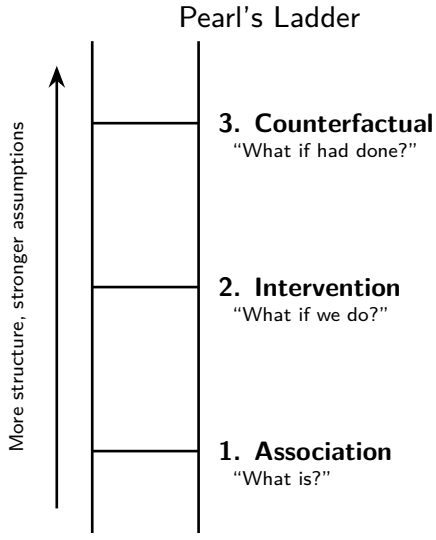
Variables per machine-shift:

- Machine group (Small / Medium / Large)
- Shift (First / Second)
- Number of changeovers
- Running hours in the shift
- **Breakdown** (yes/no)

Your model flags a machine. Now what?

1. **Disagreement** — maintenance and operations disagree on the fix
2. **High stakes** — you can't A/B test a production line
3. **Confounders everywhere** — bigger machines break more *and* get more changeovers

A framework: Pearl's Ladder of Causation

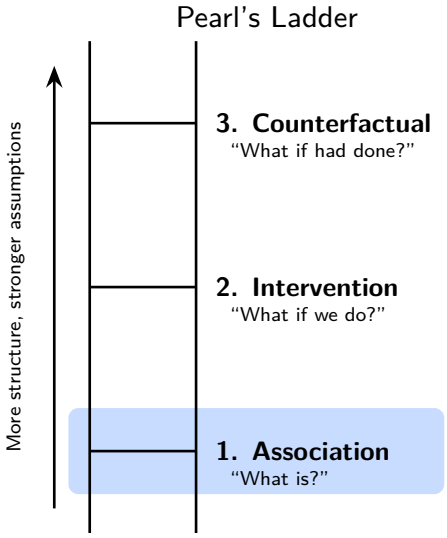


Each rung adds structure — and power:

- **Rung 1 — Association:** Add a DAG (your assumptions).
  Get testable predictions.
- **Rung 2 — Intervention:** Interpret edges as causal.
  Estimate what happens if you *act*.
- **Rung 3 — Counterfactual:** Add structural equations.
  Ask "what if we had done differently?"

Pearl & Mackenzie, *The Book of Why* (2018)

Rung 1: Association — "What is?"



Observe patterns and correlations

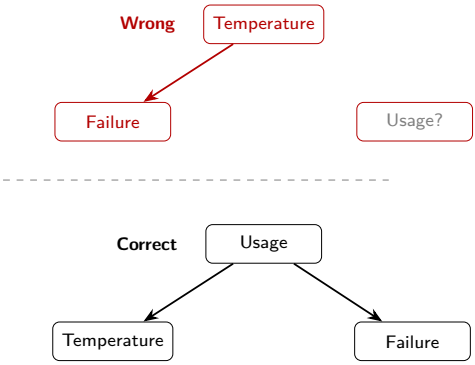
- $P(Y|X)$ — probability given we *observe*

ADD: A DAG encoding your assumptions

GET: Testable implications

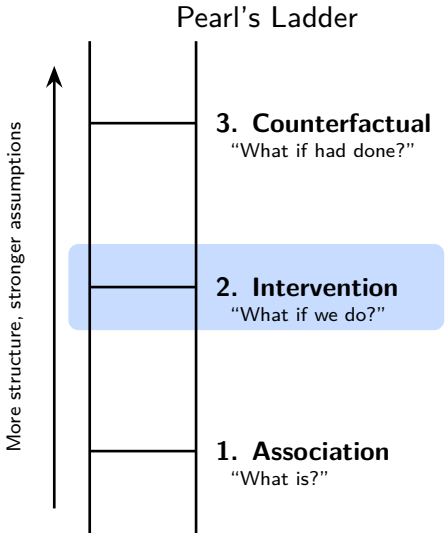
- Different DAGs → different conditional independencies
- The DAG is a *falsifiable* hypothesis

Rung 1: The DAG is a *falsifiable* hypothesis




```
impliedConditionalIndependencies('
  dag {Usage -> Temperature;
  Usage -> Failure}')
# Falr _||_ Tmpr | Usag
```

Rung 2: Intervention — "What if we do?"



Imagine (or perform) interventions

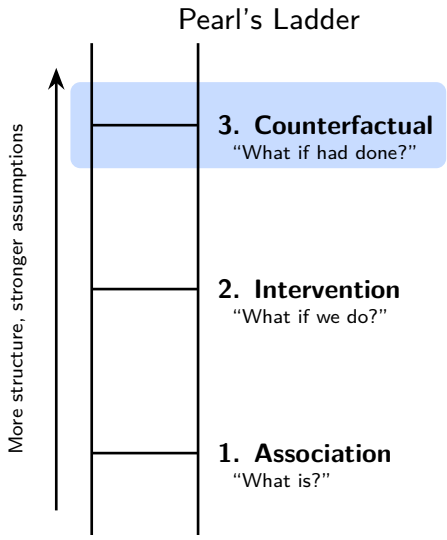
- $P(Y|\text{do}(X))$ — probability if we *force* X
- **do**$(X) \neq$ **observe** X

ADD: Interpret edges as *causal*

GET: Causal effect estimation

- "Graph surgery" — predict the result of an RCT from observational data
- No need to actually run the experiment

Rung 3: Counterfactual — "What if we had done?"



Specific machines, alternative histories

"What if *this machine* had been on Second Shift?"

ADD: Structural equations (functional forms)

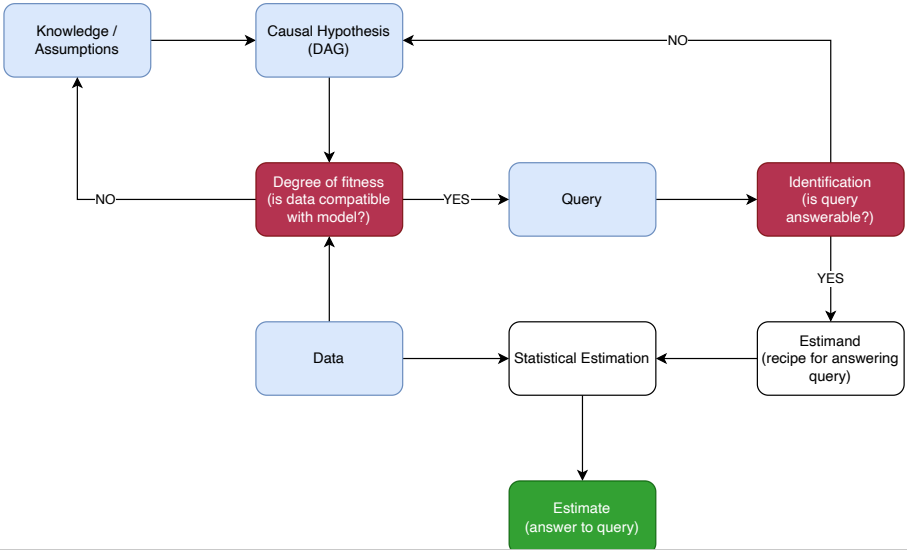
GET:

- Individual-level root cause attribution
- After-action review, blame assessment

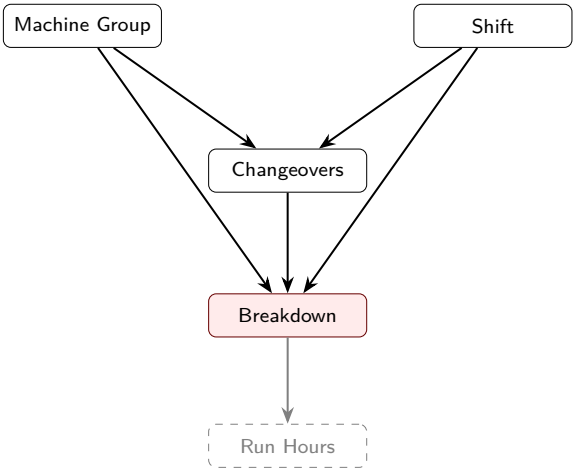
Hardest rung — strongest assumptions.

Focus today: Rungs 1 & 2

Pearl's Ladder in action: the causal inference workflow



Step 1: Draw the DAG (with the engineers)



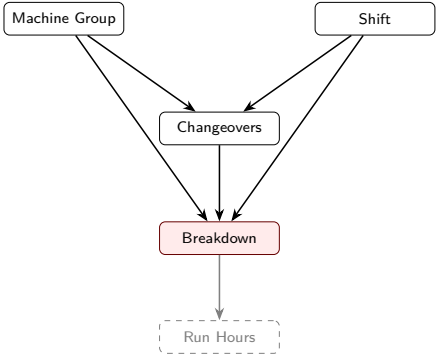
1. **Confounder:** Machine group causes both changeovers and breakdowns
2. **Mediator:** Changeovers lie on the path from shift to breakdown
3. **Descendant:** Run hours is *caused by* breakdown — do **not** condition on it

The DAG = your engineering knowledge.

Step 2: Test the DAG against data

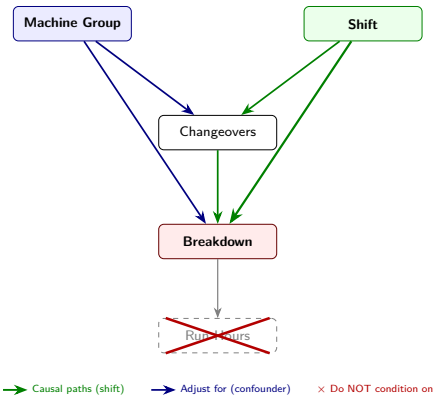
```
factory_dag <- dagitty('dag {
  machine_group -> breakdown
  machine_group -> changeovers
  shift         -> breakdown
  shift         -> changeovers
  changeovers   -> breakdown
  breakdown     -> run_hours }')

impliedConditionalIndependencies(
  factory_dag)
# chng _||_ rn_h | brkd
# mch_ _||_ rn_h | brkd
# mch_ _||_ shft
# rn_h _||_ shft | brkd
```



The DAG predicts that **run hours** is independent of everything else *once you condition on breakdown*.

Step 3: Identify — the backdoor criterion



Causal question: How strong is the direct relationship between shift and different breakdown rates?
(First Shift = day shift, Second Shift = night shift)

```
adjustmentSets(factory_dag,
  exposure = "shift",
  outcome  = "breakdown",
  effect   = "direct")
# { changeovers, machine_group }
```

Direct effect: adjust for machine group and changeovers.

A worked example: the full walkthrough

Everything we've shown so far — and more — is in a detailed walkthrough:

- Structure learning with `bnlearn`
- DAG testing with `dagitty`
- Backdoor adjustment
- Collider bias demonstration
- Heterogeneous treatment effects with `grf`

`https://theclarkeorbit.github.io/interrogating-your-twin-a-causal-infere html`

Step 4: Estimate

Model (logistic)	**Shift coeff**	p	**Key covariate**
Total effect (shift only)	-0.75	0.010	—
Total effect (+ machine group)	-0.75	0.011	precision gain, not bias correction
Direct effect (+ changeovers, mg)	-0.76	0.010	changeovers: -0.10, p=0.69
Biased (+ run_hours)	-0.79	0.008	run_hours: -0.16, p<0.001

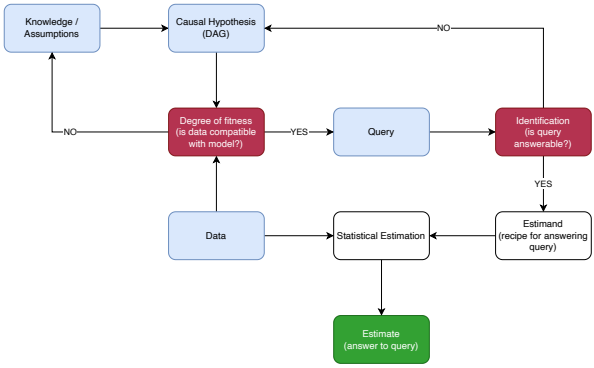
Causal ML (`grf::causal_forest`):

Machine group	**CATE**	**$/machine**
Medium	+3.3 pp	$1,649
Large	+3.0 pp	$1,484
Small	+2.5 pp	$1,266

Causal model: **reassign shifts**, **target Medium machines** first.

What we just saw: the workflow in four steps

1. Knowledge → DAG
2. DAG → Testable predictions
3. Query → Adjustment set
4. Data → Estimate



Explicit. Auditable. Defensible.

Scaling up: where causal inference fits

Your existing stack:

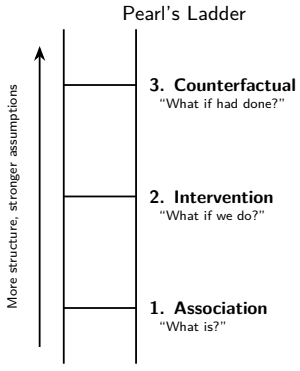
1. Sensors & SCADA — data collection
2. Data lake / historian — storage
3. ML models — pattern recognition
4. Dashboards — visualization

Causal inference plugs in as:

- A **reasoning layer** on top of ML
- DAG = your **engineering knowledge**, formalized
- Digital twin + causal model = **interventional sandbox**

Take home message

1. **Pearl's Ladder:** Seeing → Doing → Imagining
2. **DAGs** = explicit, auditable, defensible causal assumptions
3. Your predictive model tells you *what*; causal inference tells you *why* and *what to do*



We'd love to chat about your hard causal (or other) problems!

Questions?

prasanna@romulan.ltd

My LinkedIn.