Innateness in machines: Causal Inference

A talk by Riccardo Bassani & Otto Mättas

Outline

Innateness

"Innateness, AlphaZero and Artificial Intelligence", G. Marcus

- What is innateness?
- Nativism vs Empiricism
- Innateness in machines
- Innateness in Alpha Zero
- What should be innate? ➜ Causality

Causal Inference

"The Seven Tools of Causal Inference, with Reflections on Machine Learning", J. Pearl

- What is causal inference?
- Three layers of inference
- Tools for analysing causality
- Frame problem
- A possible solution for finding intelligence?

What is innateness?

The idea of something being acquired independently of learning.

Gary Marcus

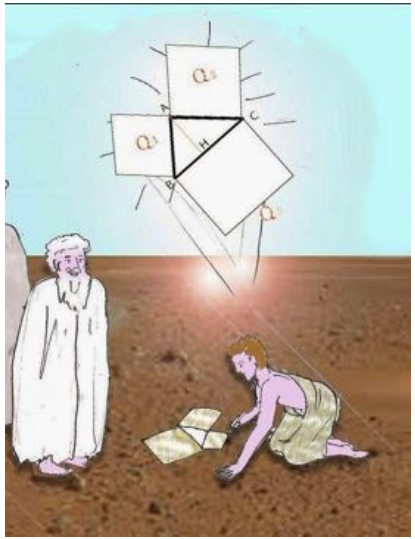
Nativism

Most basic skills are hard-wired in the brain at birth

Plato

Meno

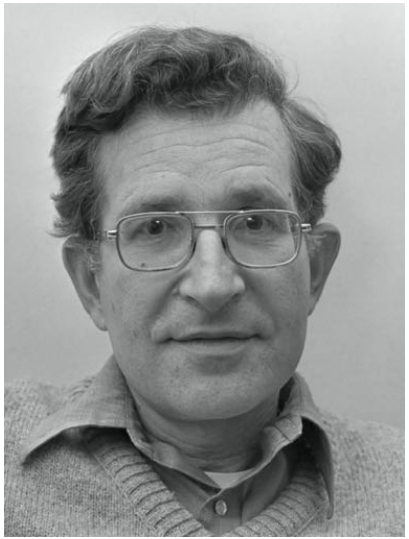
(~385 B.C.)



Chomsky

Language Acquisition Device

(1965)



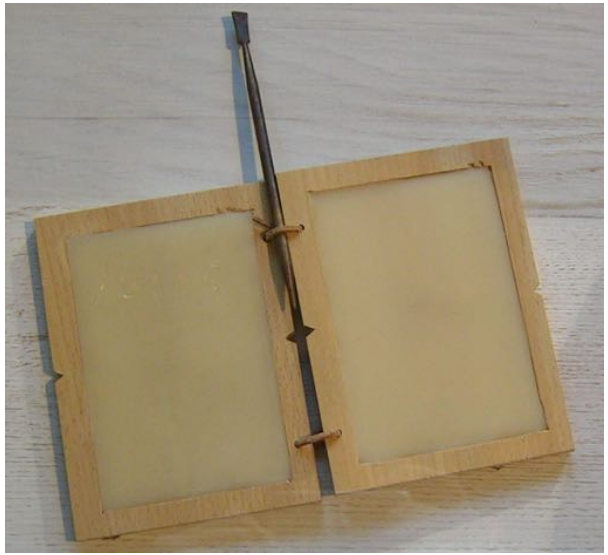
Empiricism

All knowledge is based on experience

Locke

Tabula rasa

(1694)



Innateness in machines

$$cognition = f(a, r, k, e)$$

a = innate algorithms

r = innate representational formats

k = innate knowledge

e = experience

Nativists

Empiricists

Innateness in machines

a = innate algorithms
r = innate representational formats
k = innate knowledge
e = experience

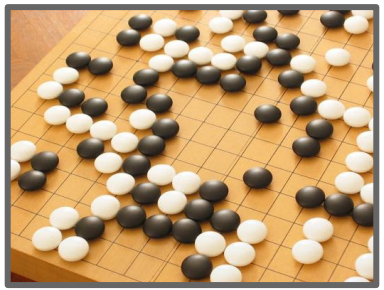
1) Evidence for innate machinery in humans
2) Proofs that ***a*** and ***r*** cannot be zero in machines

↓ still

"...most ML people believe that methods for incorporating prior knowledge in the form of symbolic rules (or their probabilistic equivalent) are too heavy-handed."

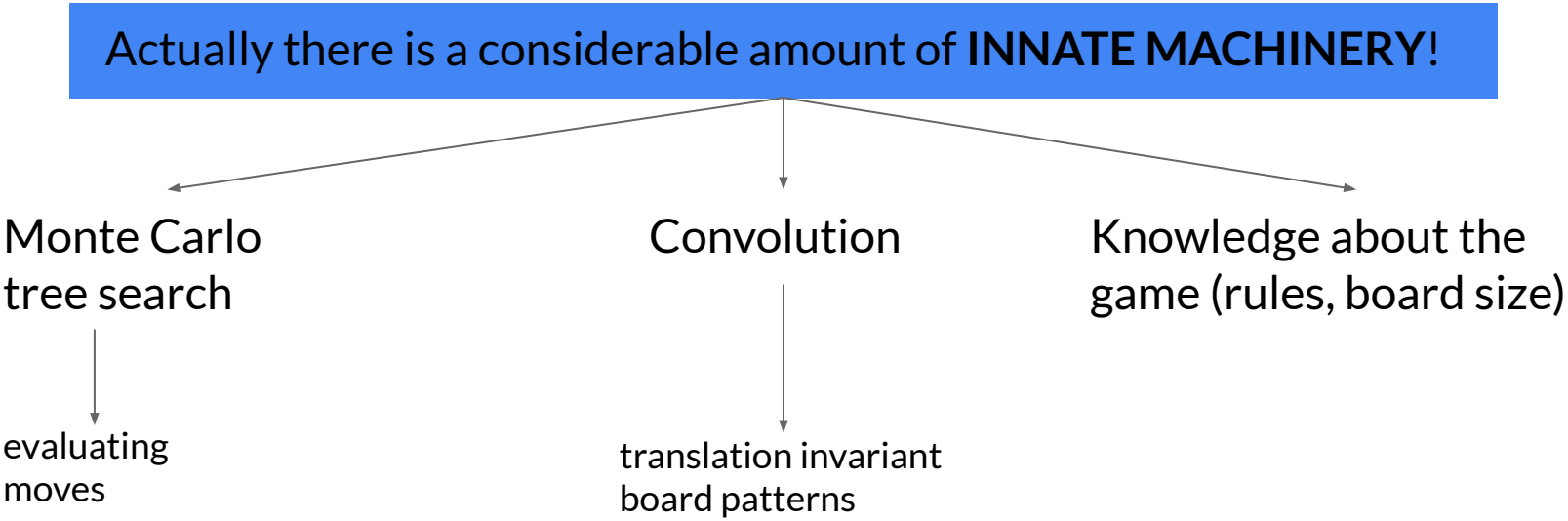
Thomas G. Dietterich

WHY? → **Alpha Zero**



Innateness in Alpha Zero

- “Mastering the game of Go without human knowledge”
- “A pure reinforcement learning approach is fully feasible, even in the most challenging of domains”



Innateness in Alpha Zero

Alpha Zero has actually extra innate machinery w.r.t.:

- Atari system (e.g. no Monte Carlo search, no game rules)
- Human mind (arguably no Monte Carlo search, certainly no game rules)

Negative argument: Alpha Zero did not renounce to innate machinery

Positive Argument: **Successful thanks to the right innate machinery**

What should be innate in an intelligent machine?

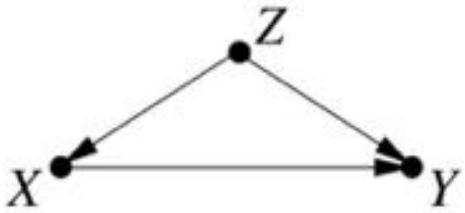
What should be innate?

10 primitives:

e.g.

- Spatiotemporal contiguity
- Capacity for cost-benefit analysis

- **Causality**

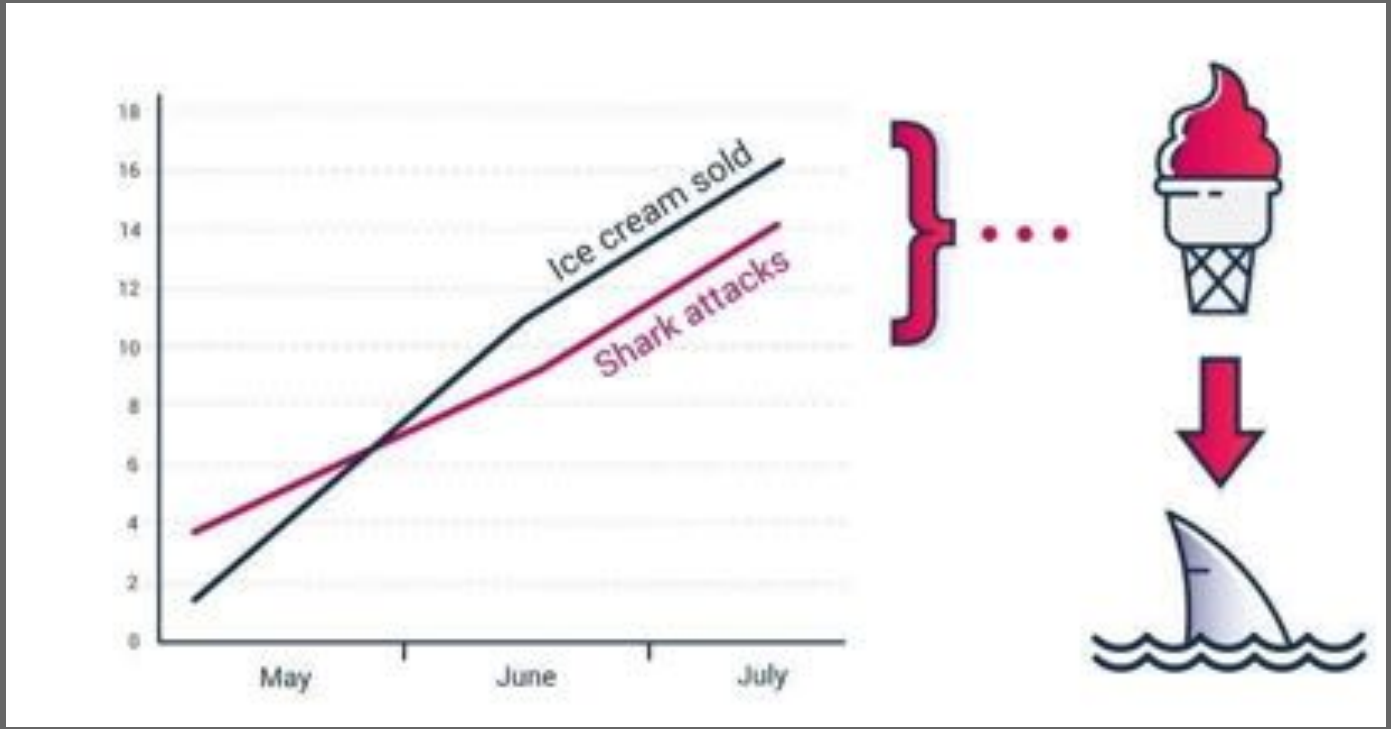


What is causal inference?

A process of analysing the response of the effect variable when the cause is changed - or more commonly, the causal relations.

Judea Pearl

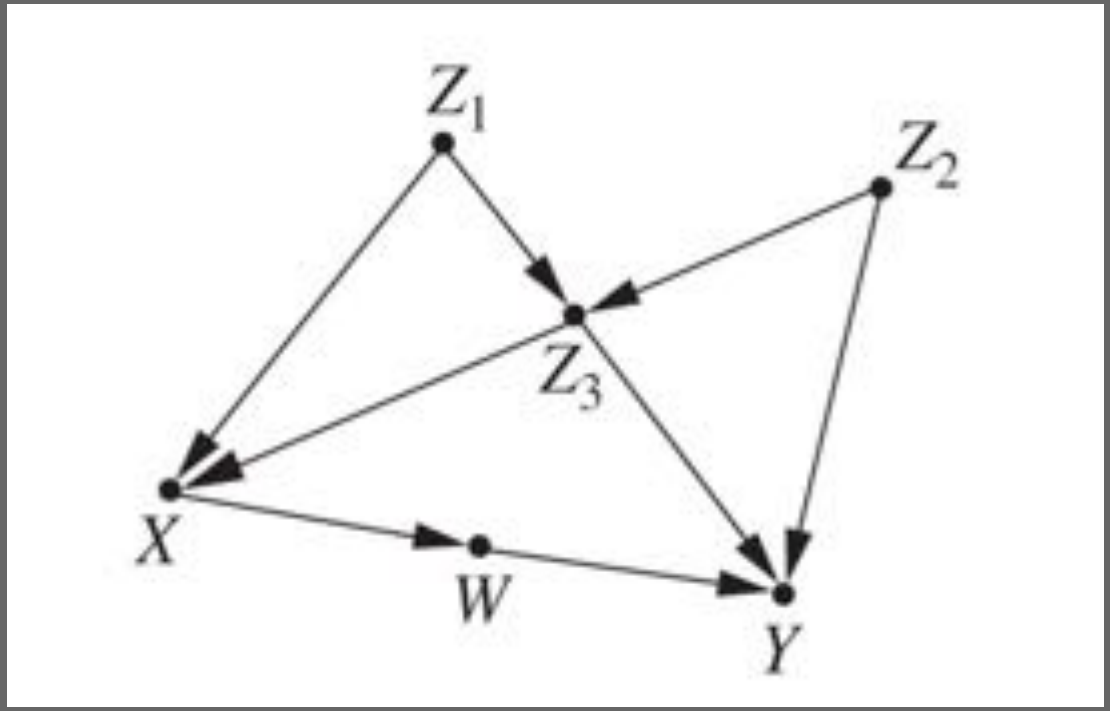
Correlation ≠ Causation



Three layers of inference

Layer (Symbol)	Typical Activity	Typical Questions	Examples
Association $P(y\|x)$	Seeing	What is? How could seeing X change my belief in Y?	What does a symptom tell me about a disease? What does a survey tell us about the election results?
Intervention $P(y\|do(x),z)$	Doing, Intervening	What if? What if I do X?	What if I take aspirin, will my headache be cured? What if we ban cigarettes?
Counterfactuals $P(y_x\|x',y')$	Imagining, Retrospecting	Why? Was it X that caused Y? What if I had acted differently?	Was it the aspirin that stopped my headache? Would Kennedy be alive had Oswald not shot him? What if I had not been smoking the past two years?

SCM: Structural Causal Model



Causality

Seven Tools of the SCM framework

- Encoding causal assumptions: Transparency and testability
- Do-calculus and the control of confounding
- The algorithmization of counterfactuals
- Mediation analysis and the assessment of direct and indirect effects
- Adaptability, external validity, and sample selection bias

- **Causal discovery**

- **Recovering from missing data**

Discussion

The Frame Problem

and recovering relationships from missing data

Discussion

Innate Machinery ⟶ Machines just do “what they are programmed for”

Discussion

Innate Machinery → ~~Machines just do "what they are programmed for"~~

Machine code

Human code

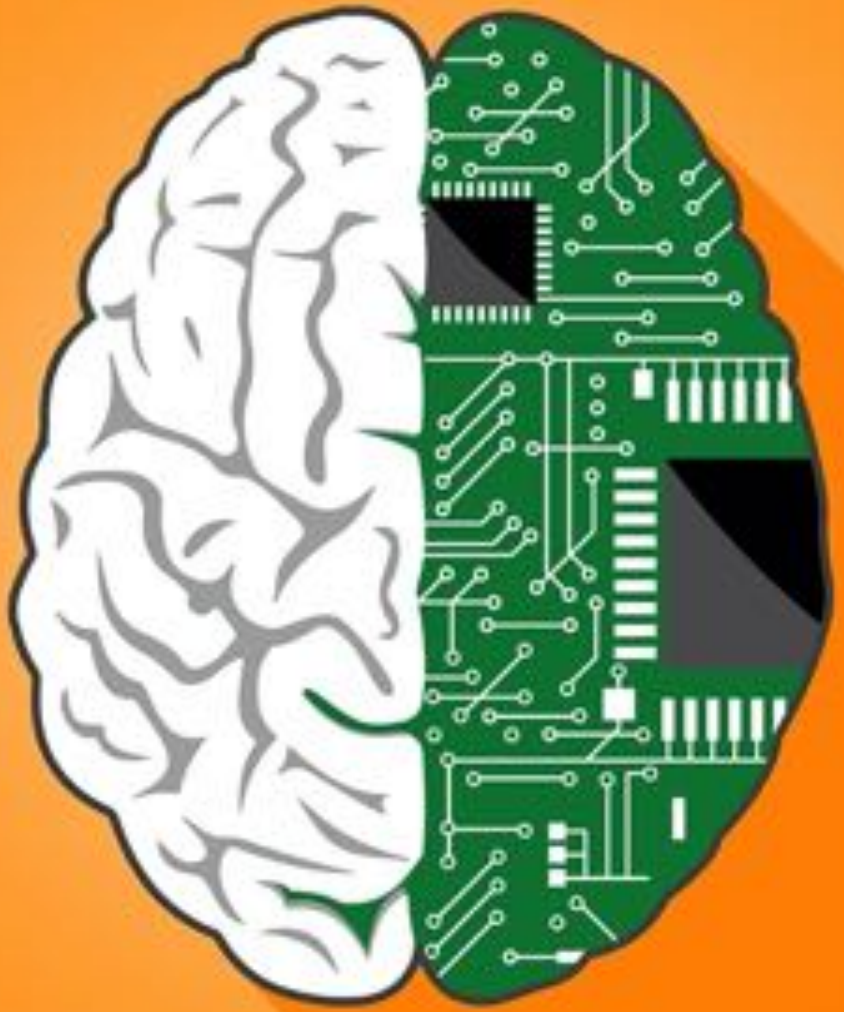
Discussion

Innateness + Learning

↓

Intelligence

What to learn?



What pre-wired?