CAP 6635 – Artificial Intelligence

Lecture 14: How do machines learn? (Part 6)

Oge Marques, PhD
Professor
College of Engineering and Computer Science
College of Business

FLORIDA ATLANTIC UNIVERSITY

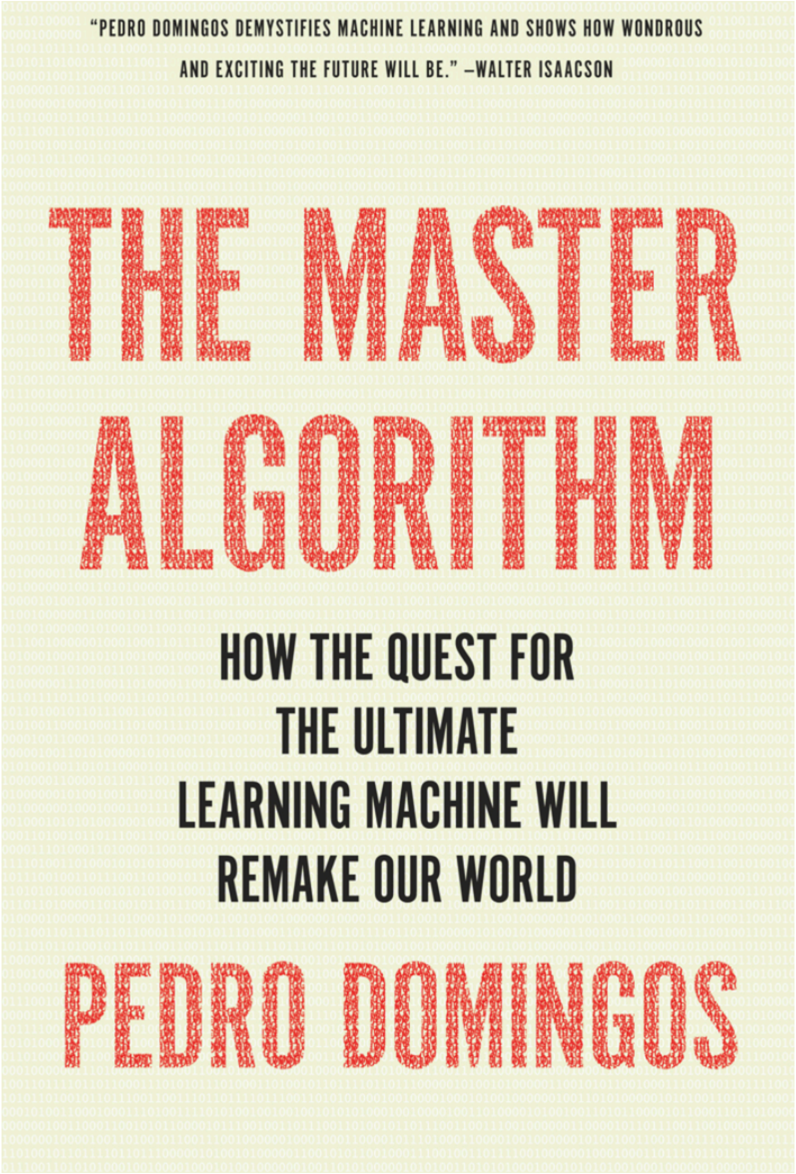
@ProfessorOge

ProfessorOgeMarques

The Master Algorithm

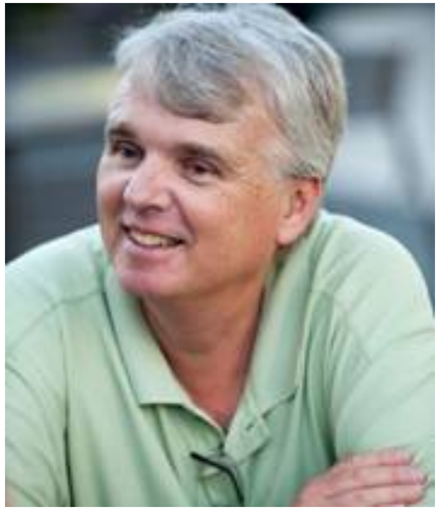
(by Pedro Domingos)

"All knowledge—past, present, and future—can be derived from data by a single, universal learning algorithm."



The Bayesians

Bayesians

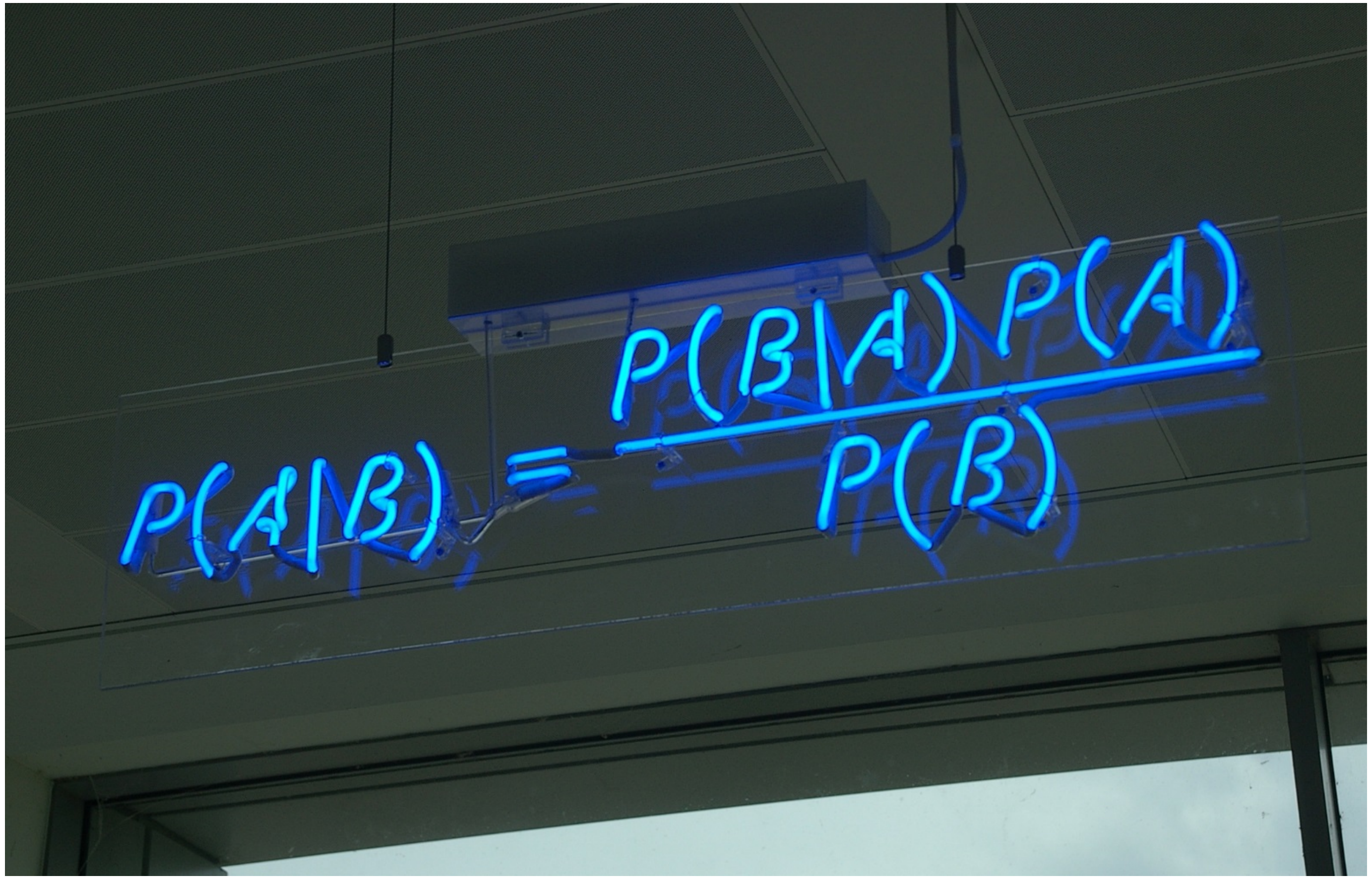


David Heckerman Judea Pearl Michael Jordan

Source: P. Domingos

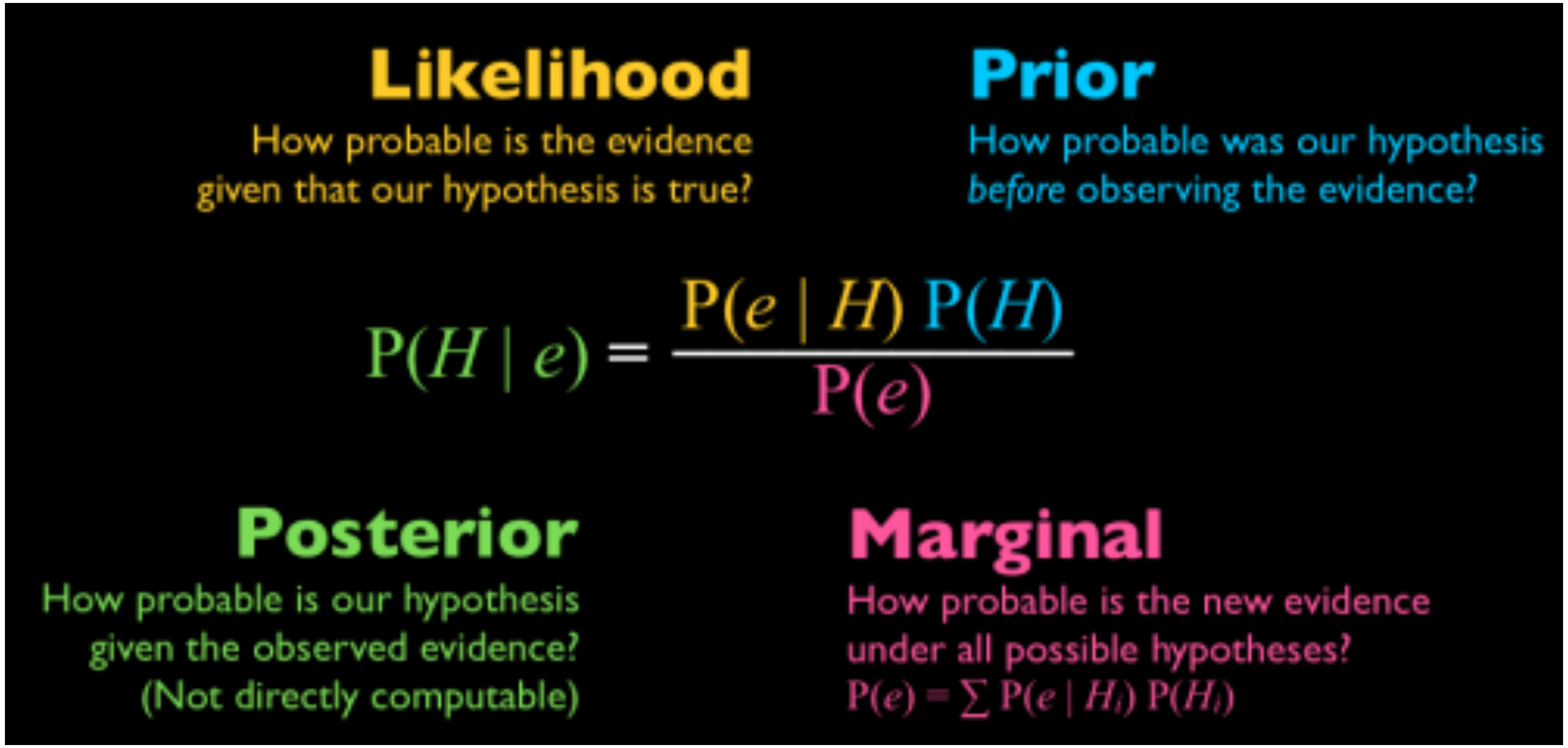
Probabilistic Inference

$$P(cause \mid effect) = P(cause) \times P(effect \mid cause) \,/\, P(effect).$$



Source: P. Domingos

Probabilistic Inference



Source: P. Domingos

Probabilistic inference: remarks

- Humans, it turns out, are not very good at Bayesian inference, at least when verbal reasoning is involved. The problem is that we tend to neglect the cause's prior probability.
- Bayes' theorem is useful because what we usually know is the probability of the effects given the causes, but what we want to know is the probability of the causes given the effects.
- Bayes' theorem as a foundation for statistics and machine learning is bedeviled not just by computational difficulty but also by extreme controversy.

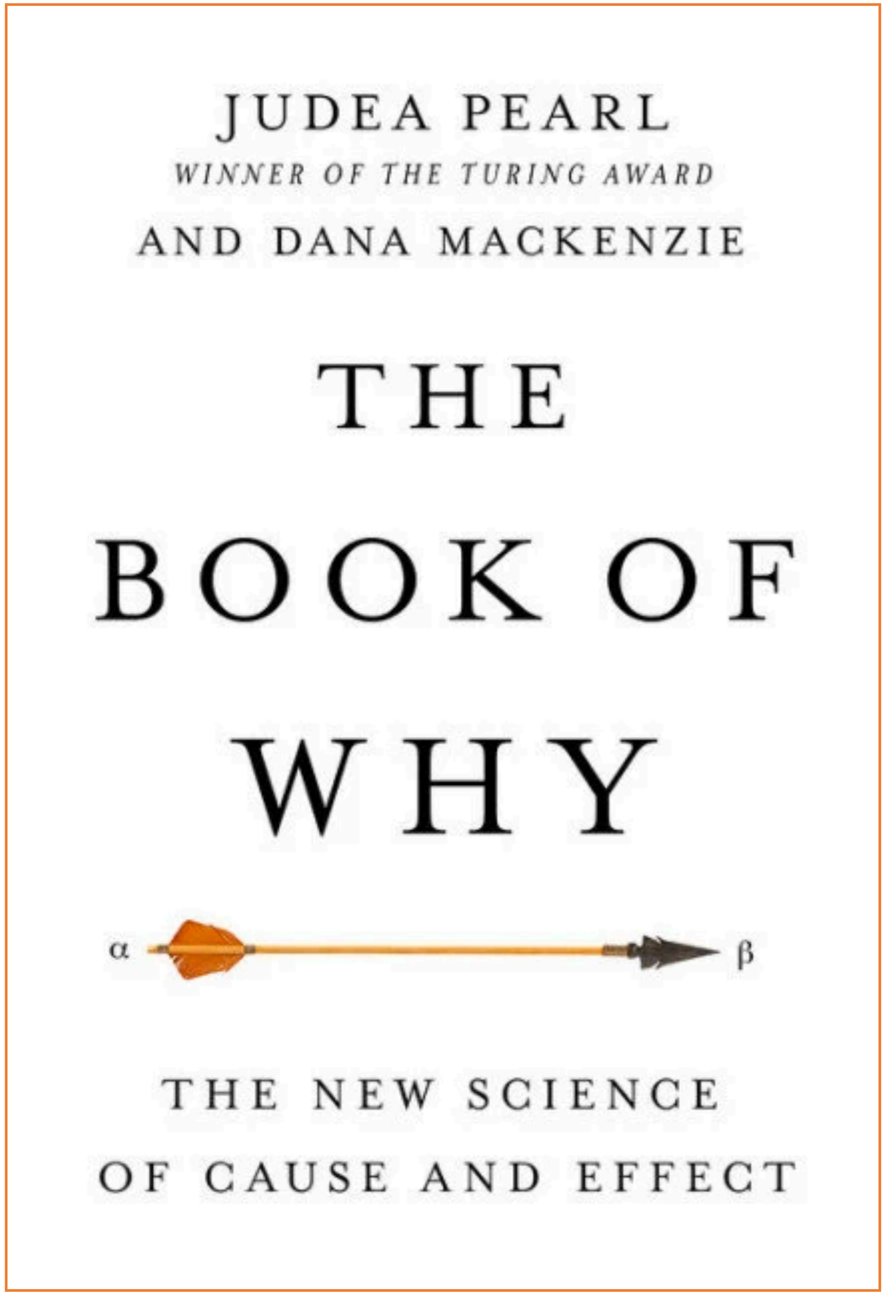
Probabilistic inference: remarks

- A learner that uses Bayes' theorem and assumes the effects are independent given the cause is called a Naïve Bayes classifier.
- As the statistician George Box famously put it:
  "All models are wrong, but some are useful."
- Naïve Bayes is a good conceptual model of a learner to use when reading the press: it captures the pairwise correlation between each input and the output, which is often all that's needed to understand references to learning algorithms in news stories.

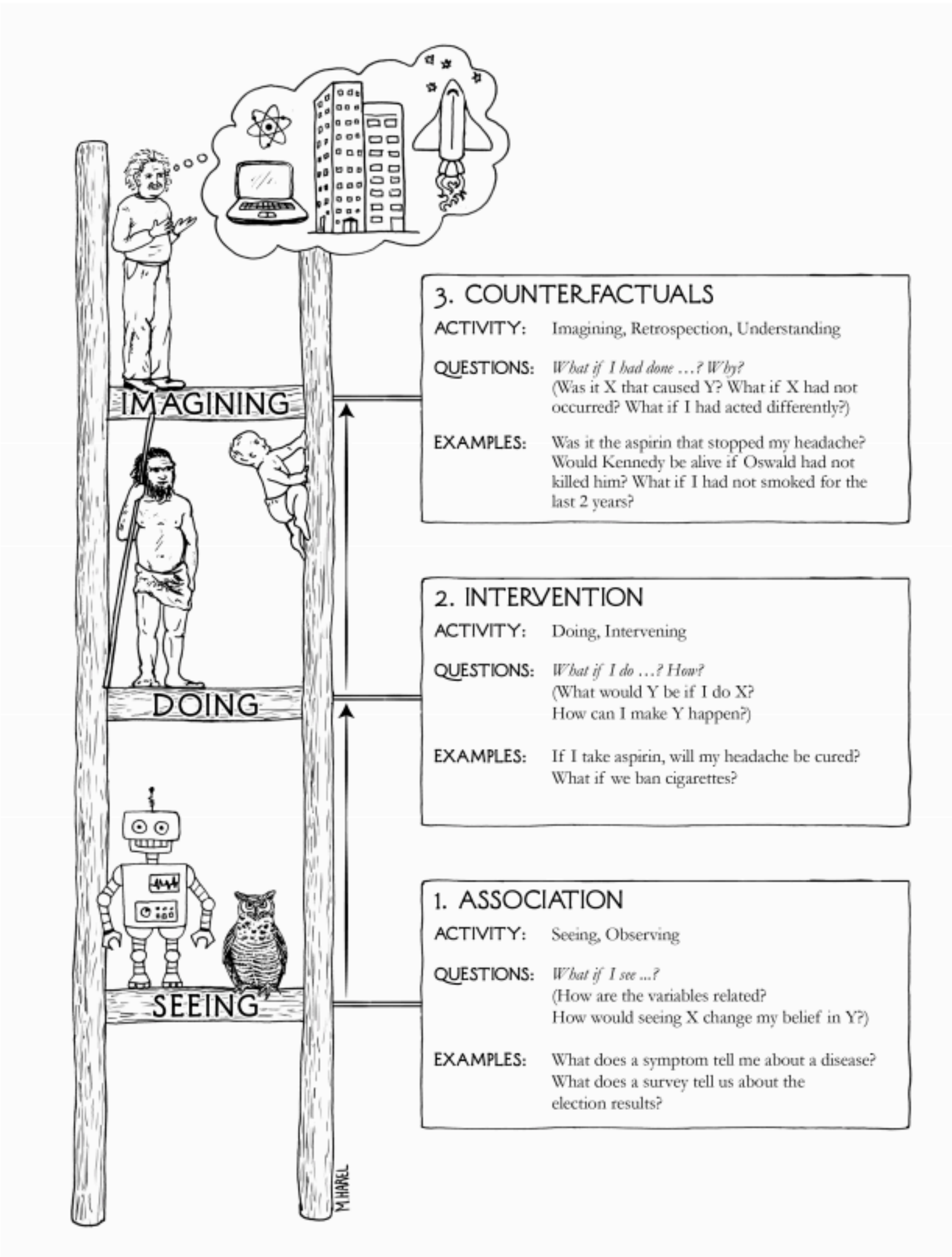
  But machine learning is not just pairwise correlations, of course, any more than the brain is just one neuron. The real action begins when we look for more complex patterns.

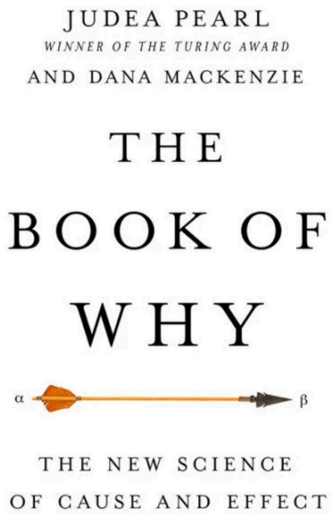
Official site: http://bayes.cs.ucla.edu/WHY/

The book of why



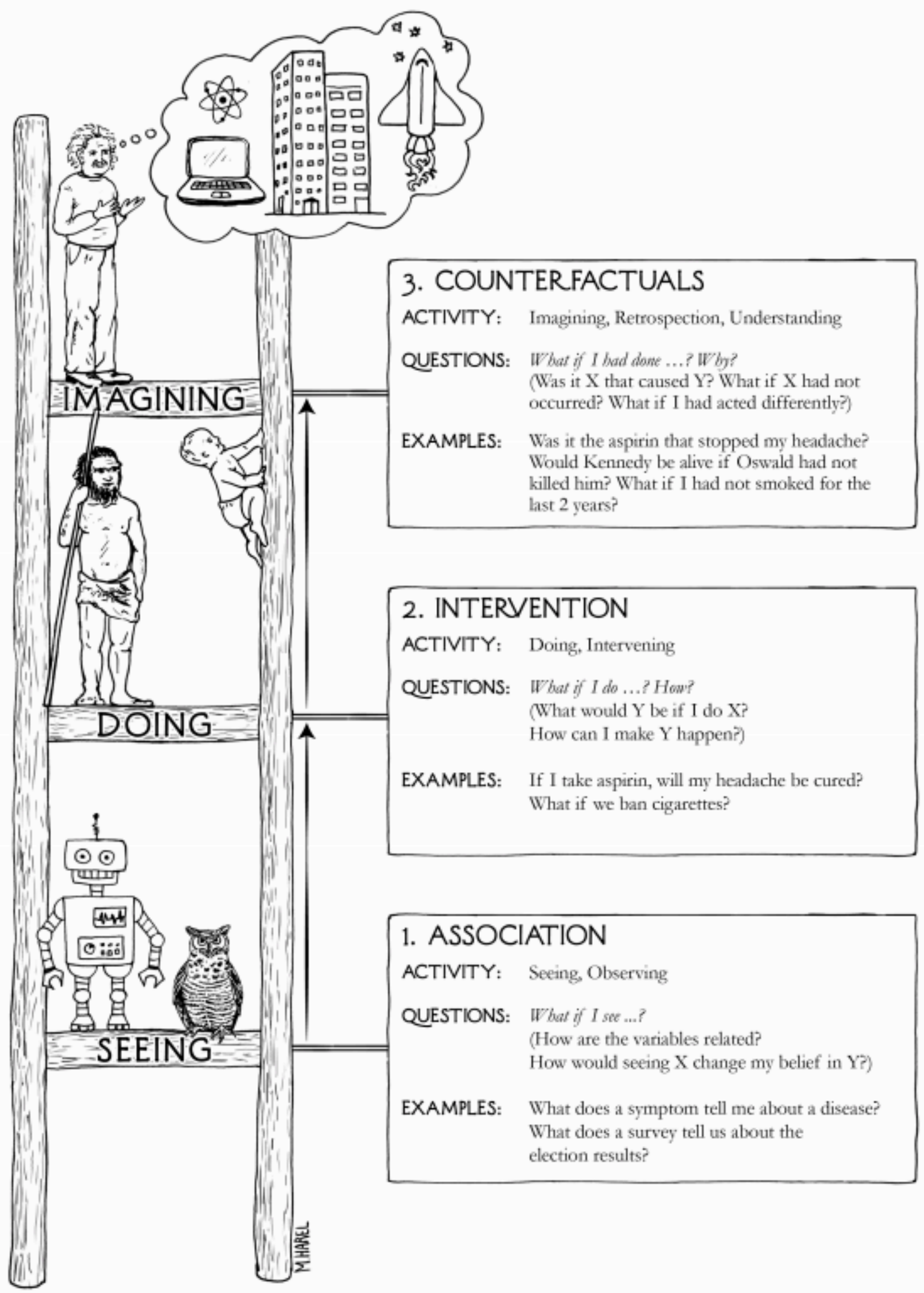
The ladder of causation





The ladder of causation

- Most animals as well as present-day learning machines are on the first rung, learning from association.
- Tool users, such as early humans, are on the second rung, if they act by planning and not merely by imitation.
- We can also use experiments to learn the effects of interventions, and presumably this is how babies acquire much of their causal knowledge.
- On the top rung, counterfactual learners can imagine worlds that do not exist and infer reasons for observed phenomena.



The ladder of causation

1. ASSOCIATION

ACTIVITY: Seeing, Observing

QUESTIONS: *What if I see ...?*
(How are the variables related?
How would seeing X change my belief in Y?)

EXAMPLES: What does a symptom tell me about a disease?
What does a survey tell us about the election results?

The ladder of causation

2. INTERVENTION

ACTIVITY: Doing, Intervening

QUESTIONS: *What if I do …? How?*
(What would Y be if I do X?
How can I make Y happen?)

EXAMPLES: If I take aspirin, will my headache be cured?
What if we ban cigarettes?

The ladder of causation

3. COUNTERFACTUALS

ACTIVITY: Imagining, Retrospection, Understanding

QUESTIONS: *What if I had done ...? Why?*
(Was it X that caused Y? What if X had not occurred? What if I had acted differently?)

EXAMPLES: Was it the aspirin that stopped my headache? Would Kennedy be alive if Oswald had not killed him? What if I had not smoked for the last 2 years?

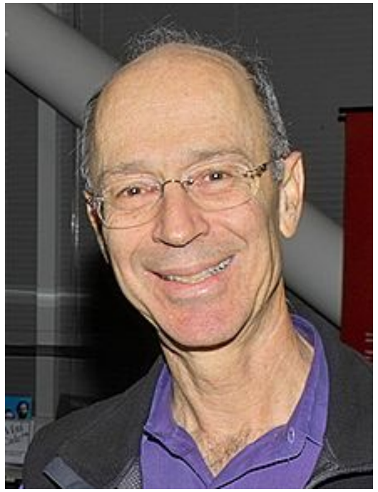
Judea Pearl: Causal Reasoning, Counterfactuals, Bayesian Networks, and the Path to AGI

- Judea Pearl is a professor at UCLA and a winner of the Turing Award, that's generally recognized as the Nobel Prize of computing.
- He is one of the seminal figures in the field of artificial intelligence, computer science, and statistics. He has developed and championed probabilistic approaches to AI, including Bayesian Networks and profound ideas in causality in general.
- These ideas are important not just for AI, but to our understanding and practice of science.
- But in the field of AI, the idea of causality, cause and effect, to many, lies at the core of what is currently missing and what must be developed in order to build truly intelligent systems.
- For this reason, and many others, his work is worth returning to often.

https://lexfridman.com/judea-pearl/

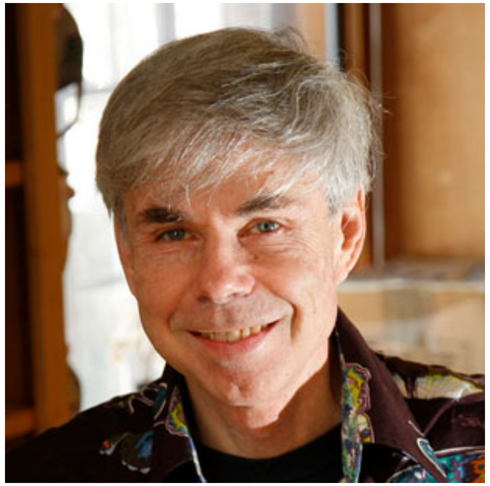
The Analogizers

Analogizers



Peter Hart

Vladimir Vapnik



Douglas Hofstadter

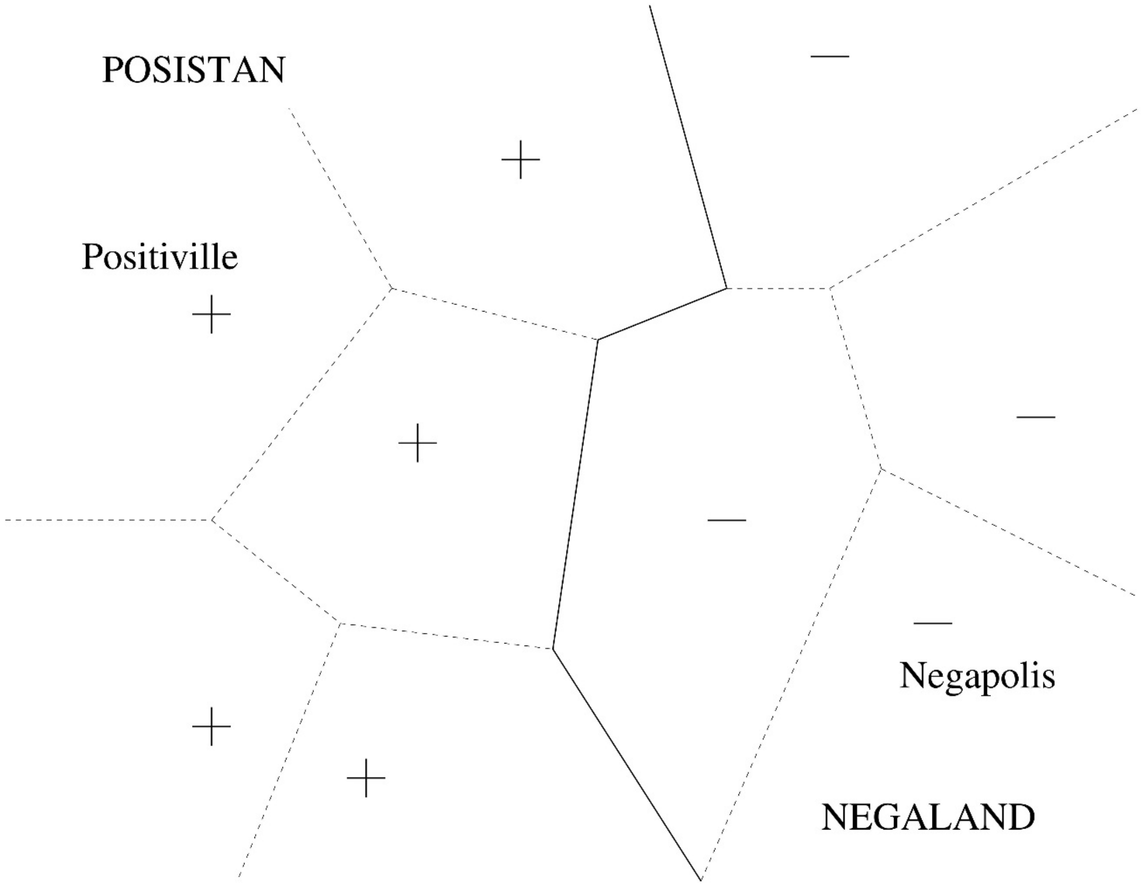
Source: P. Domingos

Vladimir Vapnik: Statistical Learning

- Vladimir Vapnik is the co-inventor of support vector machines, support vector clustering, VC theory, and many foundational ideas in statistical learning.
- His work has been cited over 170,000 times.
- He has some very interesting ideas about AI and the nature of learning, especially on the limits of our current approaches and the open problems in the field.

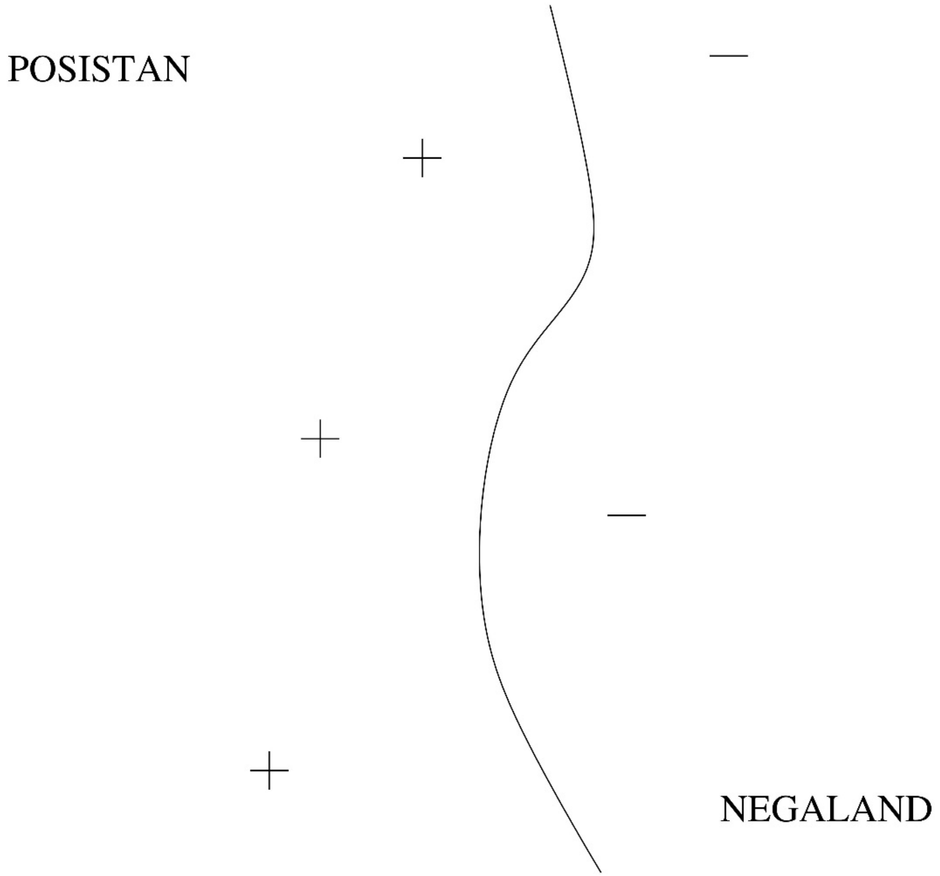
https://lexfridman.com/vladimir-vapnik/

Nearest Neighbor



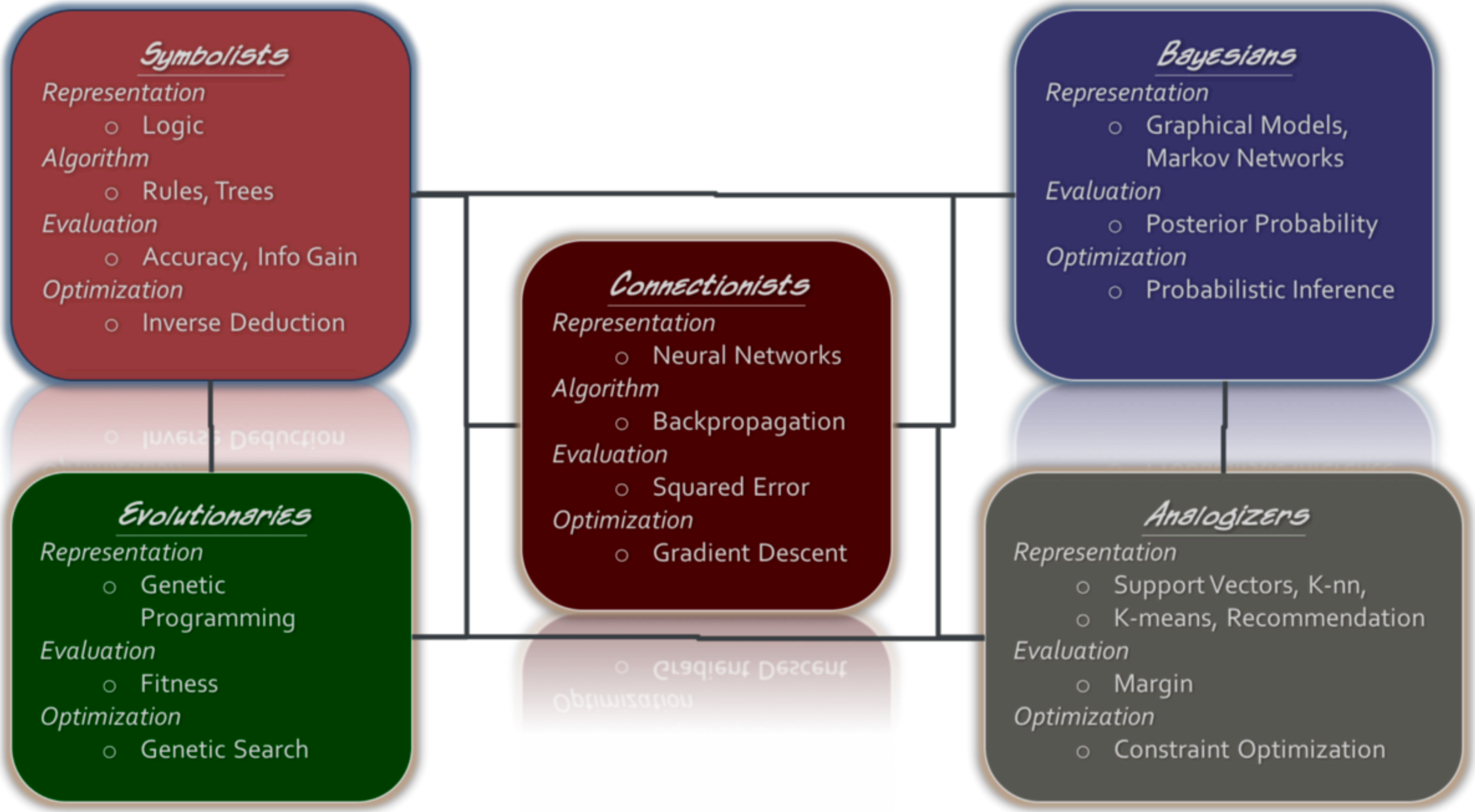
Source: P. Domingos

Kernel Machines



Source: P. Domingos

The Master Algorithm

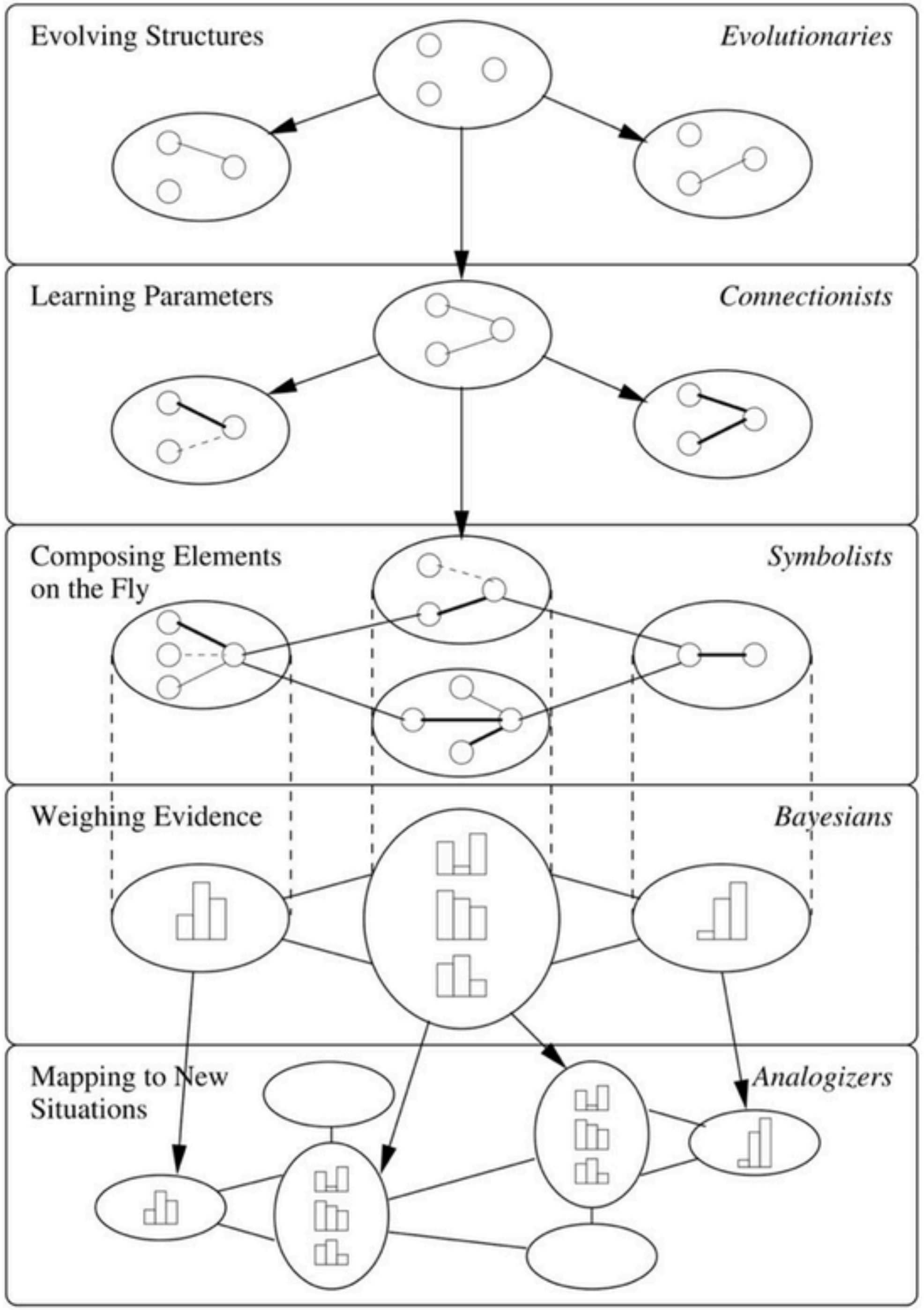

Symbolists
Representation
Logic
Algorithm
Rules, Trees
Evaluation
Accuracy, Info Gain
Optimization
Inverse Deduction
Bayesians
Representation
Graphical Models, Markov Networks
Evaluation
Posterior Probability
Optimization
Probabilistic Inference
Connectionists
Representation
Neural Networks
Algorithm
Backpropagation
Evaluation
Squared Error
Optimization
Gradient Descent
Evolutionaries
Representation
Genetic Programming
Evaluation
Fitness
Optimization
Genetic Search
Analogizers
Representation
Support Vectors, K-nn,
K-means, Recommendation
Evaluation
Margin
Optimization
Constraint Optimization

The Big Picture

Tribe	Problem	Solution
Symbolists	Knowledge composition	Inverse deduction
Connectionists	Credit assignment	Backpropagation
Evolutionaries	Structure discovery	Genetic programming
Bayesians	Uncertainty	Probabilistic inference
Analogizers	Similarity	Kernel machines

But what we really need is
a single algorithm that solves all five!

Source: P. Domingos



Source: P. Domingos

“The Master Algorithm”

- Concluding Remarks (Chapter 10)
  - “In the coming decades, machine learning will affect such a broad swath of human life that one chapter of one book cannot possibly do it justice. ”
  - Online dating
  - “Digital mirror”
  - A society of models
  - Data sharing and privacy concerns
  - The future of dating, hiring, learning, etc.
  - Several futuristic scenarios and speculations