CAP 6635 – Artificial Intelligence

Lecture 7: Foundations of AI (Part 2)

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

FAU FLORIDA ATLANTIC UNIVERSITY

@ProfessorOge

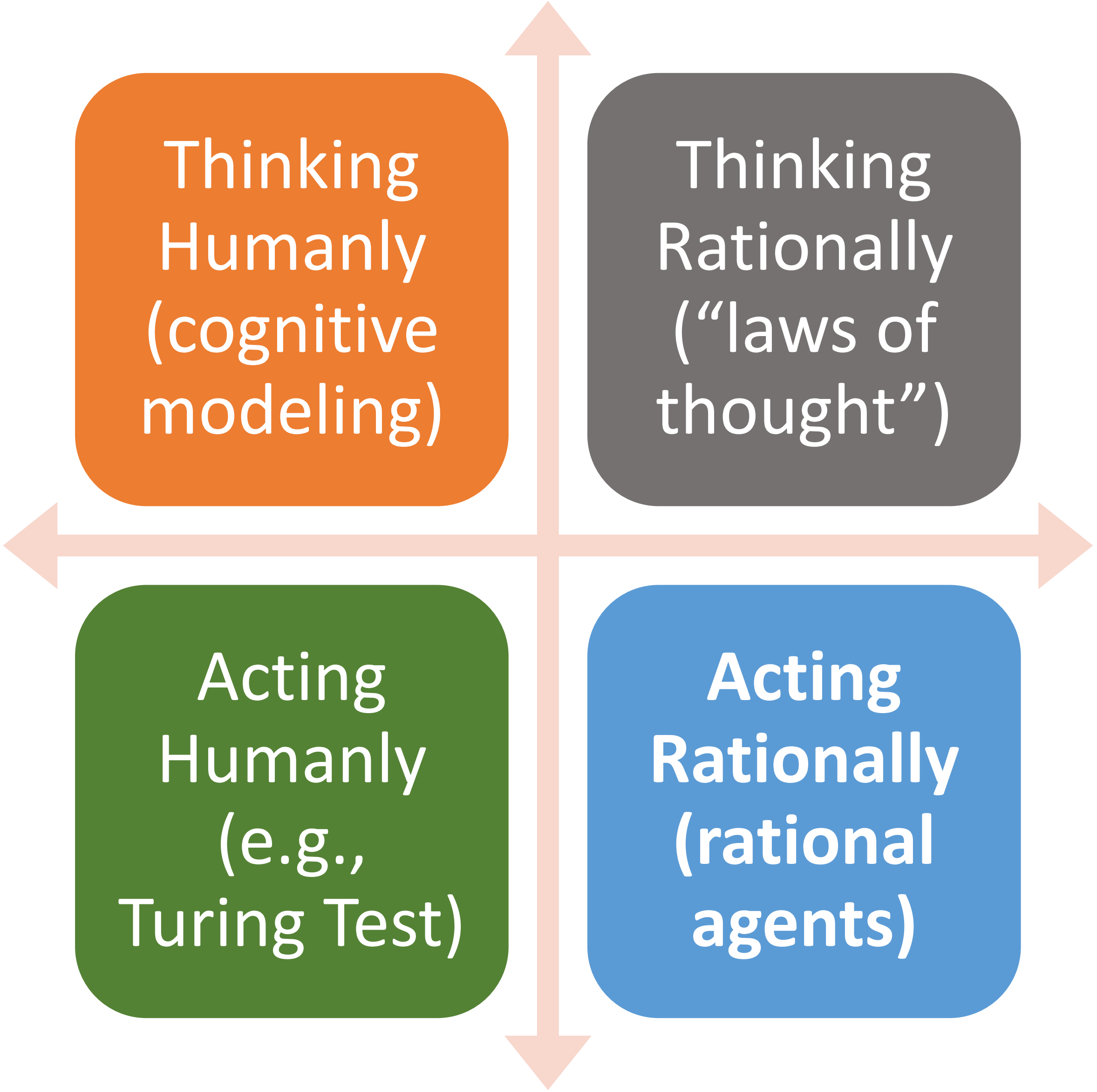
ProfessorOgeMarques

Previously on CAP 6635…

What is Artificial Intelligence?

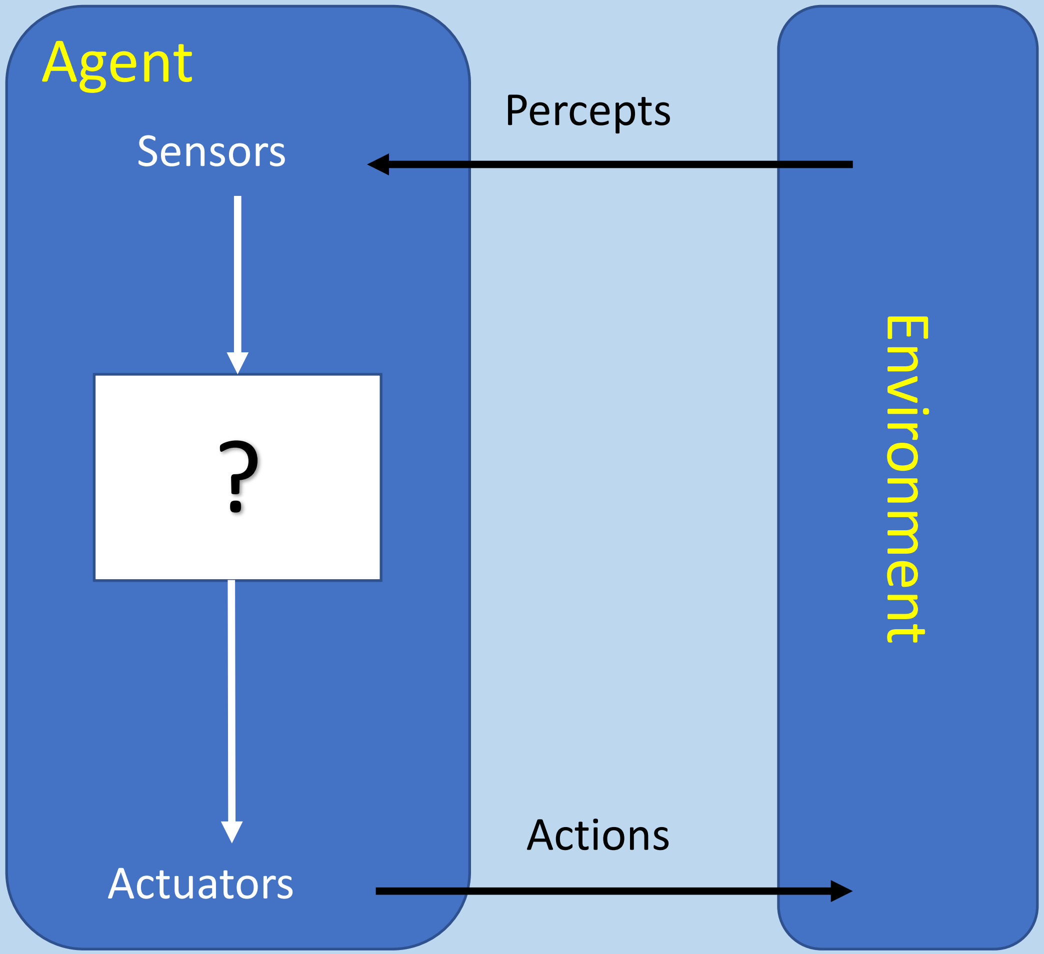
Source: Loukides and Lorica (2016) https://www.oreilly.com/ideas/what-is-artificial-intelligence

Artificial Intelligence definitions: categories



Source: (Russell & Norvig 2011)

Agents and environment



Source: (Russell & Norvig 2011)

Rational agents

- For each possible percept sequence, a rational agent should:
  - select an action that is expected to maximize its performance measure,
  - based on the evidence provided by the percept sequence and
  - whatever built-in knowledge the agent has.

Source: (Russell & Norvig 2011)

PEAS: Performance measure, Environment, Actuators, Sensors

Agent Type	Performance Measure	Environment	Actuators	Sensors
Robot soccer player	Winning game, goals for/against	Field, ball, own team, other team, own body	Devices (e.g., legs) for locomotion and kicking	Camera, touch sensors, accelerometers, orientation sensors, wheel/joint encoders
Internet book-shopping agent	Obtain re-quested/interesting books, minimize expenditure	Internet	Follow link, enter/submit data in fields, display to user	Web pages, user requests
Autonomous Mars rover	Terrain explored and reported, samples gathered and analyzed	Launch vehicle, lander, Mars	Wheels/legs, sample collection device, analysis devices, radio transmitter	Camera, touch sensors, accelerometers, orientation sensors, , wheel/joint encoders, radio receiver

Environment types

- **Fully observable** (vs. partially observable): An agent's sensors give it access to the complete state of the environment at each point in time.
- **Deterministic** (vs. stochastic): The next state of the environment is completely determined by the current state and the action executed by the agent.
- **Episodic** (vs. sequential): The agent's experience is divided into atomic "episodes" (each episode consists of the agent perceiving and then performing a single action), and the choice of action in each episode depends only on the episode itself.

Environment types

- **Static** (vs. dynamic): The environment is unchanged while an agent is deliberating. (The environment is semi-dynamic if the environment itself does not change with the passage of time, but the agent's performance score does)
- **Discrete** (vs. continuous): A limited number of distinct, clearly defined percepts and actions.
- **Single agent** (vs. multiagent): An agent operating by itself in an environment.

What are the environment types of…

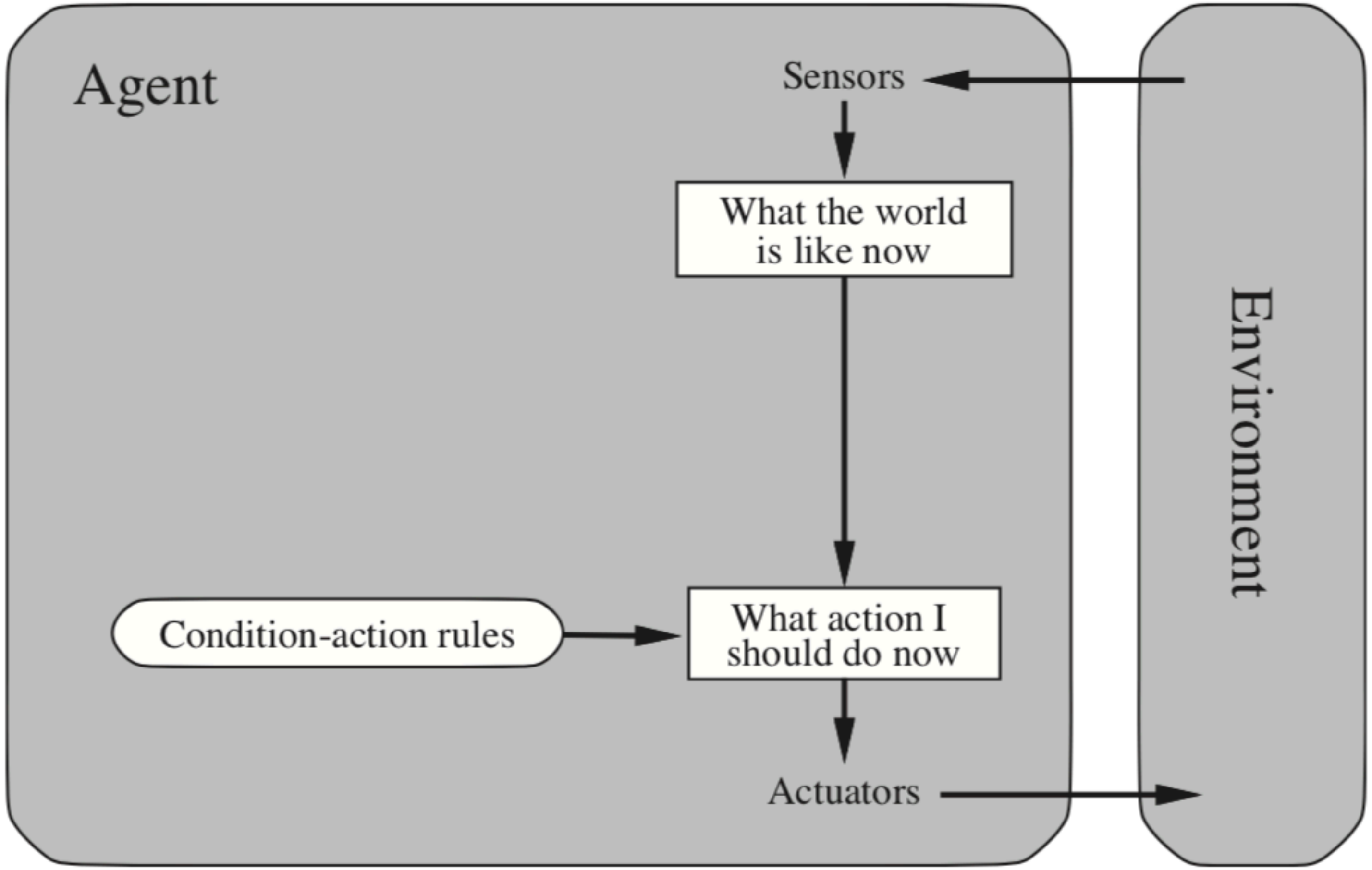
- Tic-tac-toe
- Chess
- Go
- DOTA 2
- Self-driving vehicles
- The real world

- **Fully** vs. **partially** observable
- **Deterministic** vs. **stochastic**
- **Episodic** vs. **sequential**
- **Static** vs. **dynamic**
- **Discrete** vs. **continuous**
- **Single agent** vs. **multiagent**

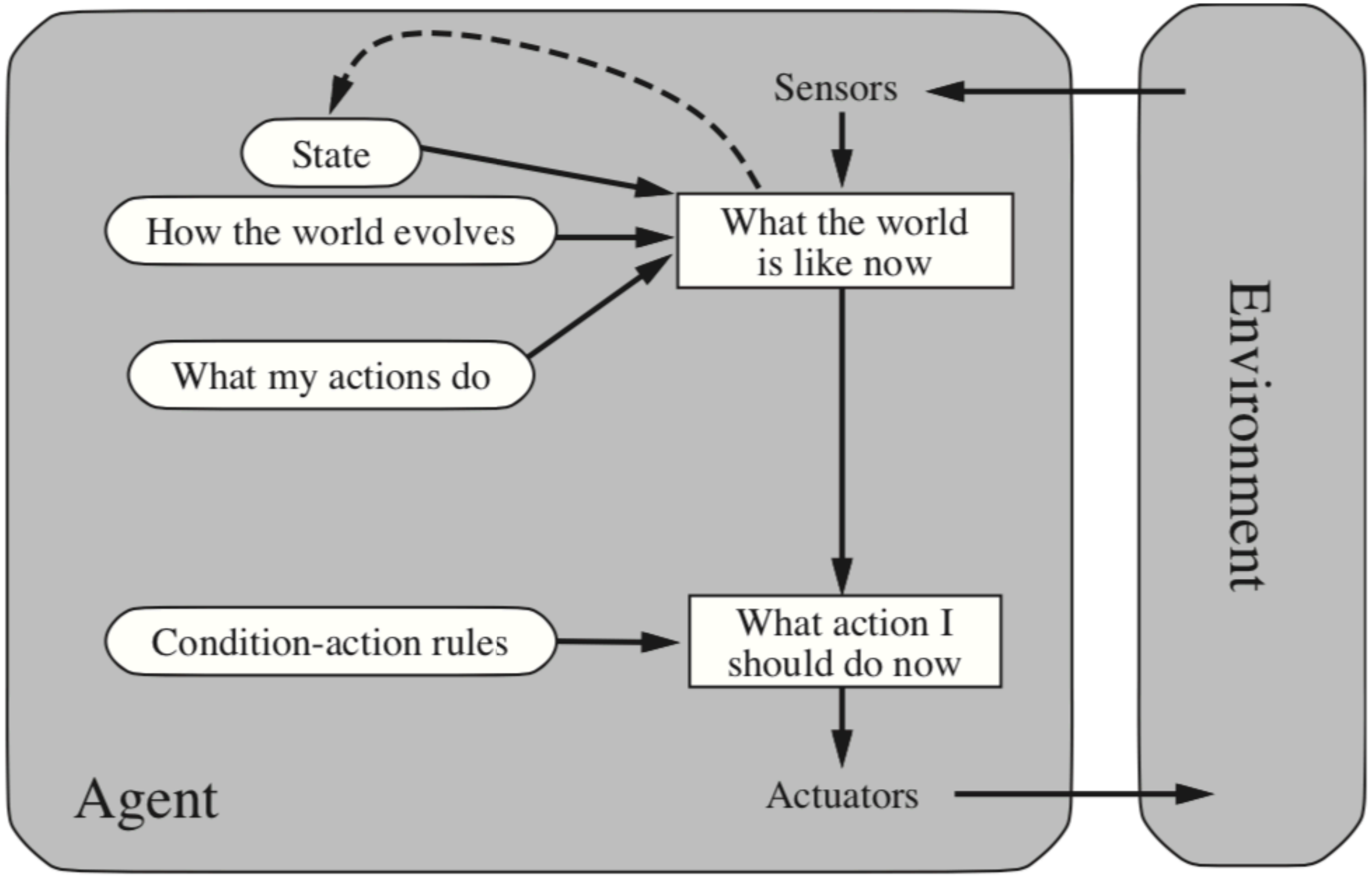
Agent types (in order of increasing generality)

- Simple reflex agents
- Agents that keep track of the world
- Goal-based agents
- Utility-based agents

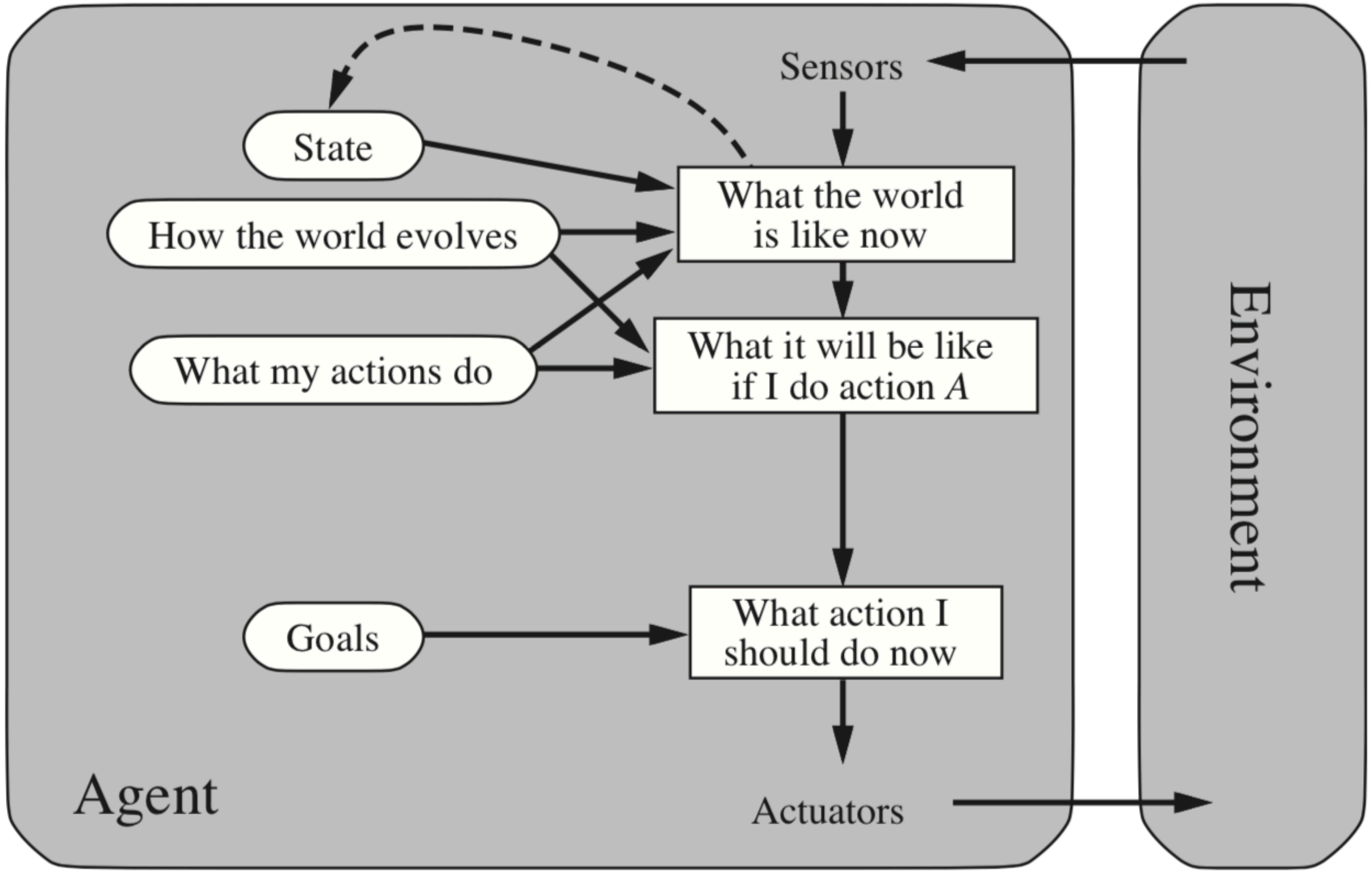
Simple reflex agents



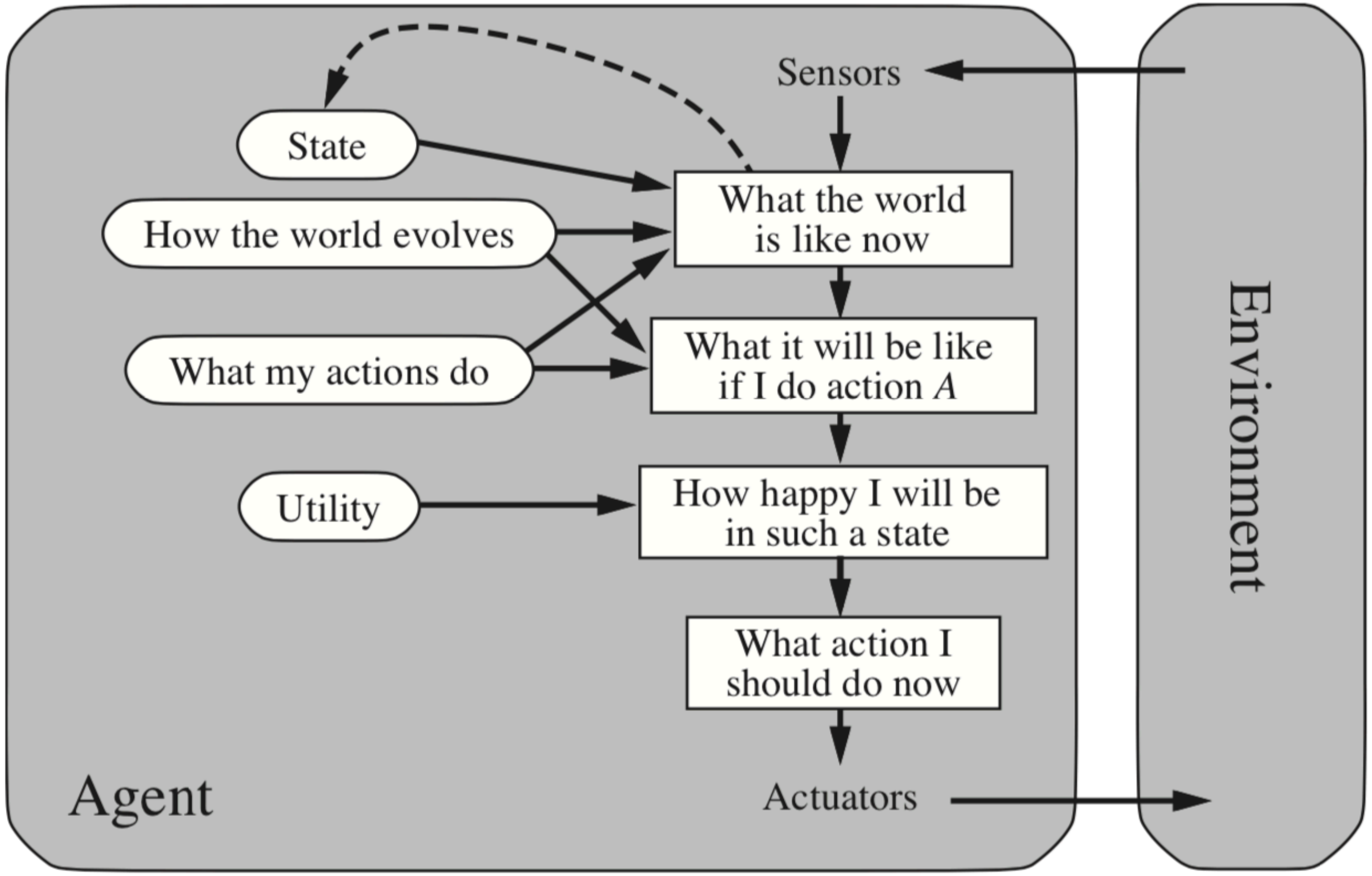
Agents that keep track of the world (Agents with internal states)



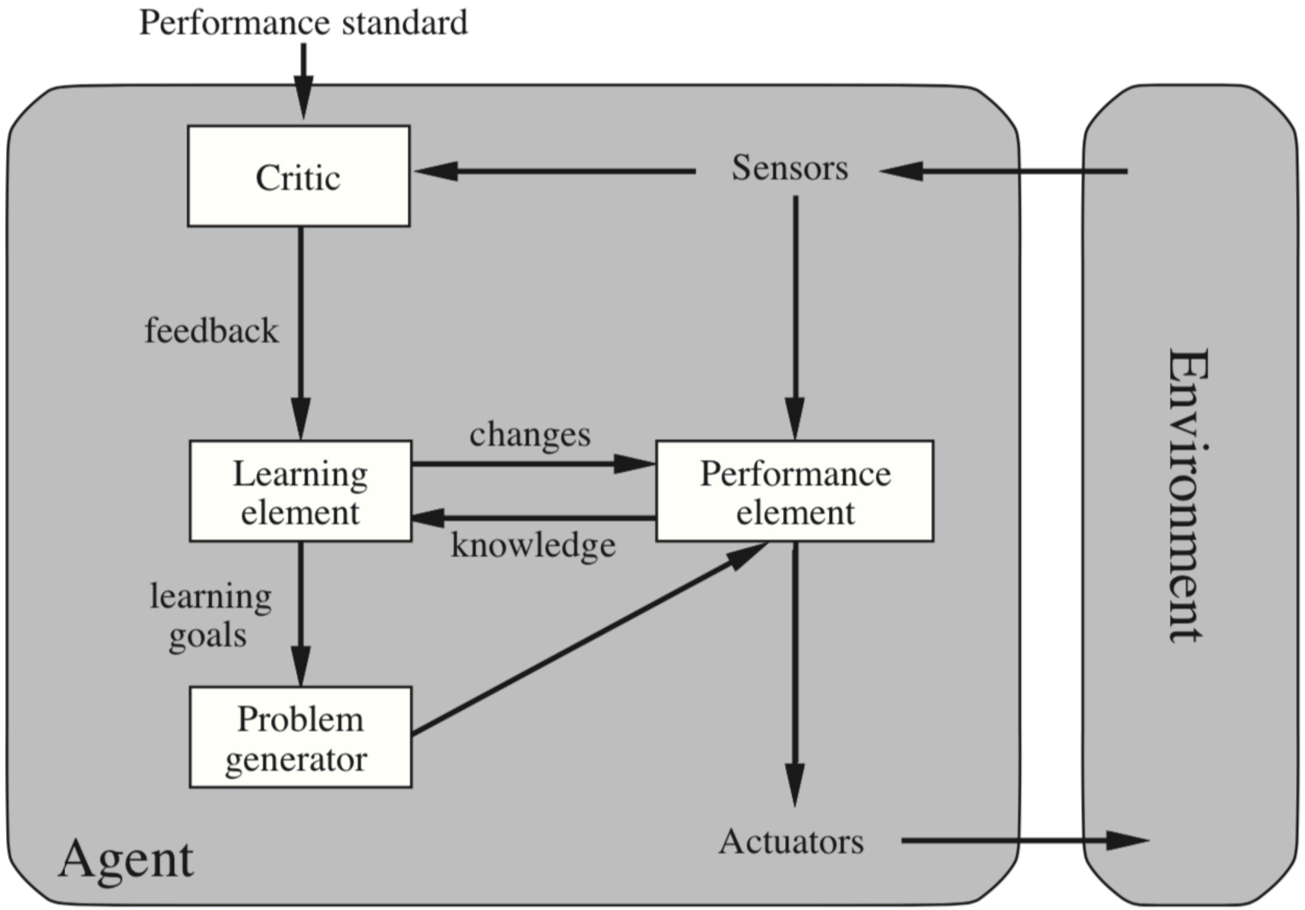
Model-based, goal-based agent



Model-based, utility-based agent



A general learning agent



Dimensions of AI

- **Strength** (how intelligent is it?)
- **Breadth** (does it solve a narrowly defined problem, or is it general?)
- **Training** (how does it learn?)
- **Capabilities** (what kinds of problems are we asking it to solve?)
- **Autonomy** (are AIs assistive technologies, or do they act on their own?)

Source: Loukides and Lorica (2016)

Narrow vs. General AI

- What we have today is mostly **narrow AI**

“You can add narrow AIs ad infinitum, but a pile of narrow intelligences will never add up to a general intelligence.”

Source: Loukides and Lorica (2016)

New material starts here…

Learning

Reasons to learn

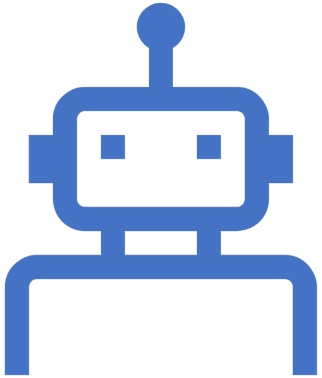
An agent is learning if it improves its performance on future tasks after making observations about the world.

Why would we want an agent to learn?

1. Designers cannot anticipate all possible situations that the agents might find themselves in.
2. Designers cannot anticipate all changes over time.
3. Sometimes human programmers have no idea how to program a solution themselves ▶ their solutions will consist of <u>building models that can learn from examples</u>.

How Do Machines Learn?

- Fill in gaps in existing knowledge
- Emulate the brain
- Simulate evolution
- Systematically reduce uncertainty
- Notice similarities between old and new



Source: P. Domingos

The 5 tribes of Machine Learning

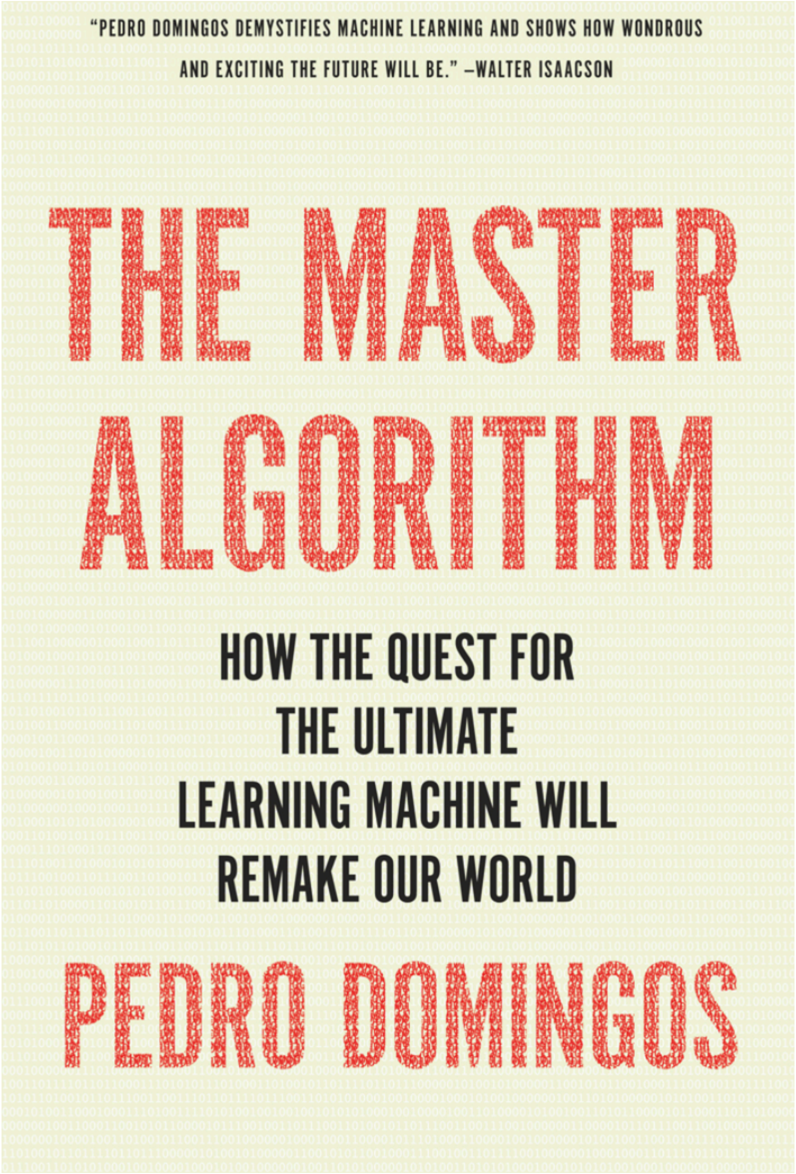
TRIBE	ORIGINS	MASTER ALGORITHM
Symbolists	Logic, philosophy	Inverse deduction
Connectionists	Neuroscience	Backpropagation
Evolutionaries	Evolutionary biology	Genetic programming
Bayesians	Statistics	Probabilistic inference
Analogizers	Psychology	Kernel machines

Source: P. Domingos

The search for a “Master Algorithm”

The Master Algorithm

(by Pedro Domingos)



A look at

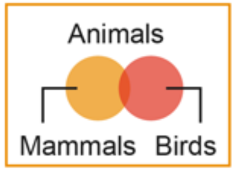
Machine learning evolution

Overview

For decades, individual “tribes” of artificial intelligence researchers have vied with one another for dominance. Is the time ripe now for tribes to collaborate? They may be forced to, as collaboration and algorithm blending are the only ways to reach true artificial general intelligence (AGI). Here’s a look back at how machine learning methods have evolved and what the future may look like.

What are the five tribes?

Symbolists



Use symbols, rules, and logic to represent knowledge and draw logical inference

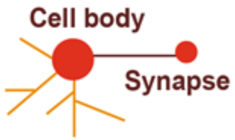
Favored algorithm
Rules and decision trees

Bayesians

Assess the likelihood of occurrence for probabilistic inference

Favored algorithm
Naive Bayes or Markov

Connectionists



Recognize and generalize patterns dynamically with matrices of probabilistic, weighted neurons

Favored algorithm
Neural networks

Evolutionaries

Generate variations and then assess the fitness of each for a given purpose

Favored algorithm
Genetic programs

Analogizers

Optimize a function in light of constraints (“going as high as you can while staying on the road”)

Favored algorithm
Support vectors

Source: Pedro Domingos, *The Master Algorithm*, 2015

Image from: http://usblogs.pwc.com/emerging-technology/machine-learning-evolution-infographic/

Phases of evolution

1980s

Predominant tribe
Symbolists

Architecture
Server or mainframe

Predominant theory
Knowledge engineering

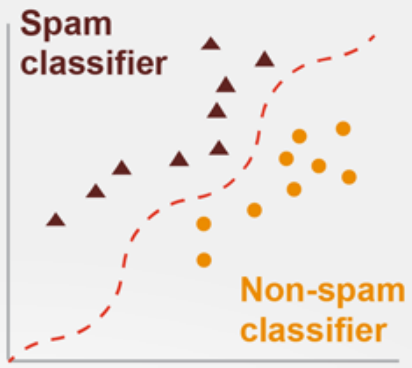
Basic decision logic:
Decision support systems with limited utility

1990s to 2000

Predominant tribe
Bayesians

Architecture
Small server clusters

Predominant theory
Probability theory



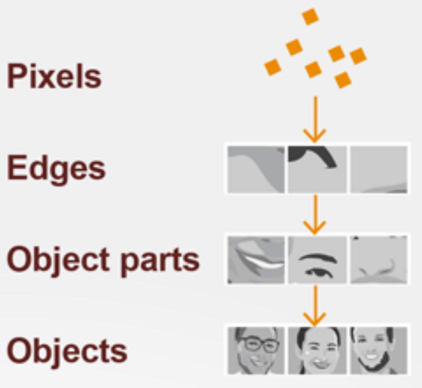
Classification:
Scalable comparison and contrast that's good enough for many purposes

Early to mid-2010s

Predominant tribe
Connectionists

Architecture
Large server farms (the cloud)

Predominant theory
Neuroscience and probability



Recognition:
More precise image and voice recognition, translation, sentiment analysis, etc.

Image from: http://usblogs.pwc.com/emerging-technology/machine-learning-evolution-infographic/

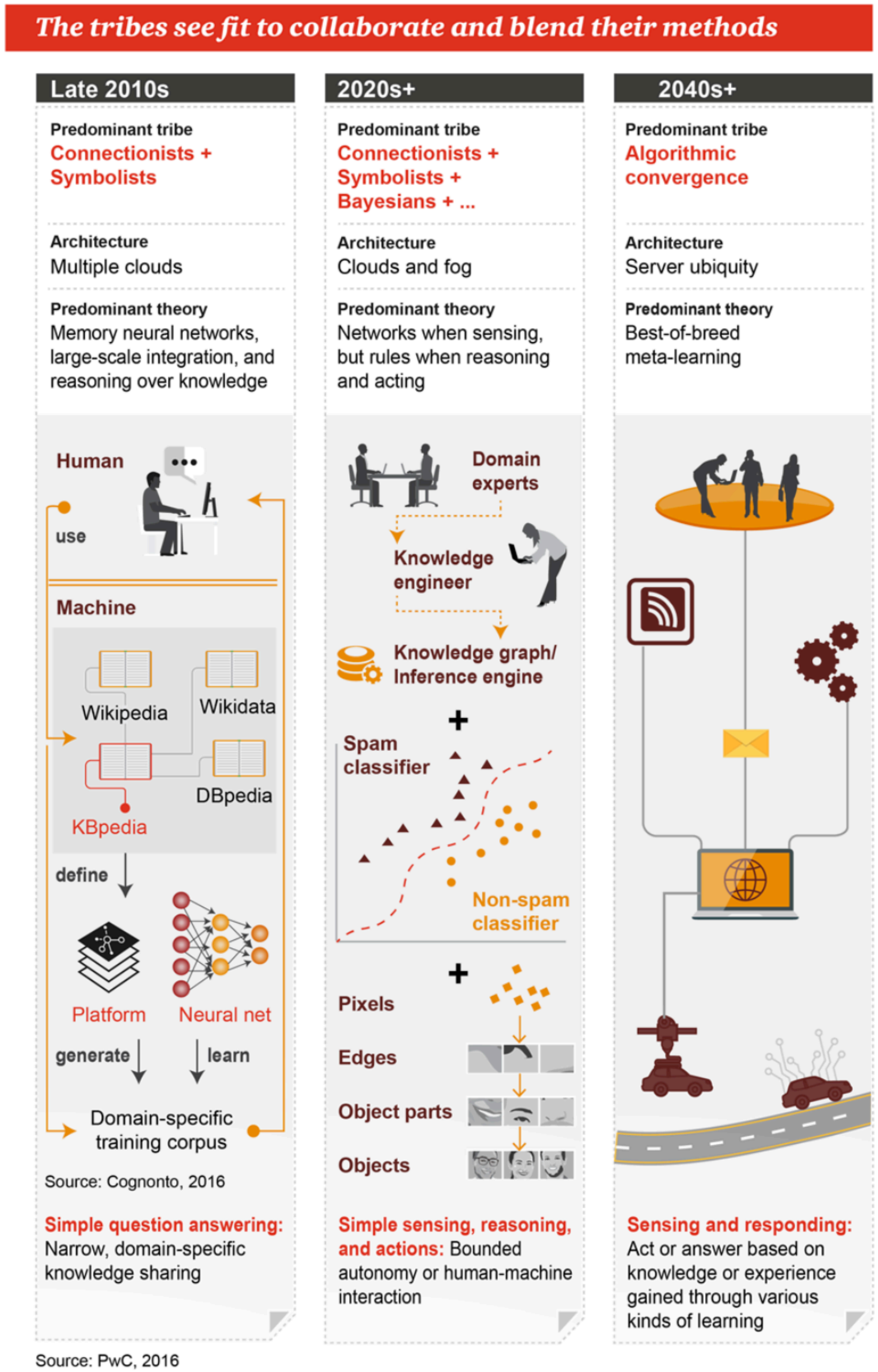


Image from: http://usblogs.pwc.com/emerging-technology/machine-learning-evolution-infographic/

The 5 tribes of Machine Learning

TRIBE	ORIGINS	MASTER ALGORITHM
Symbolists	Logic, philosophy	Inverse deduction
Connectionists	Neuroscience	Backpropagation
Evolutionaries	Evolutionary biology	Genetic programming
Bayesians	Statistics	Probabilistic inference
Analogizers	Psychology	Kernel machines

Source: P. Domingos

Symbolists

Tom Mitchell

Steve Muggleton

Ross Quinlan

Source: P. Domingos

Inverse Deduction

Addition

$$\begin{array}{r} 2 \\ +\ 2 \\ \hline =\ ? \end{array}$$

Subtraction

$$\begin{array}{r} 2 \\ +\ ? \\ \hline =\ 4 \end{array}$$

Source: P. Domingos

Inverse Deduction

Deduction

Socrates is human

\+ Humans are mortal

= ?

Induction

Socrates is human

\+ ?

= Socrates is mortal

Source: P. Domingos

Connectionists

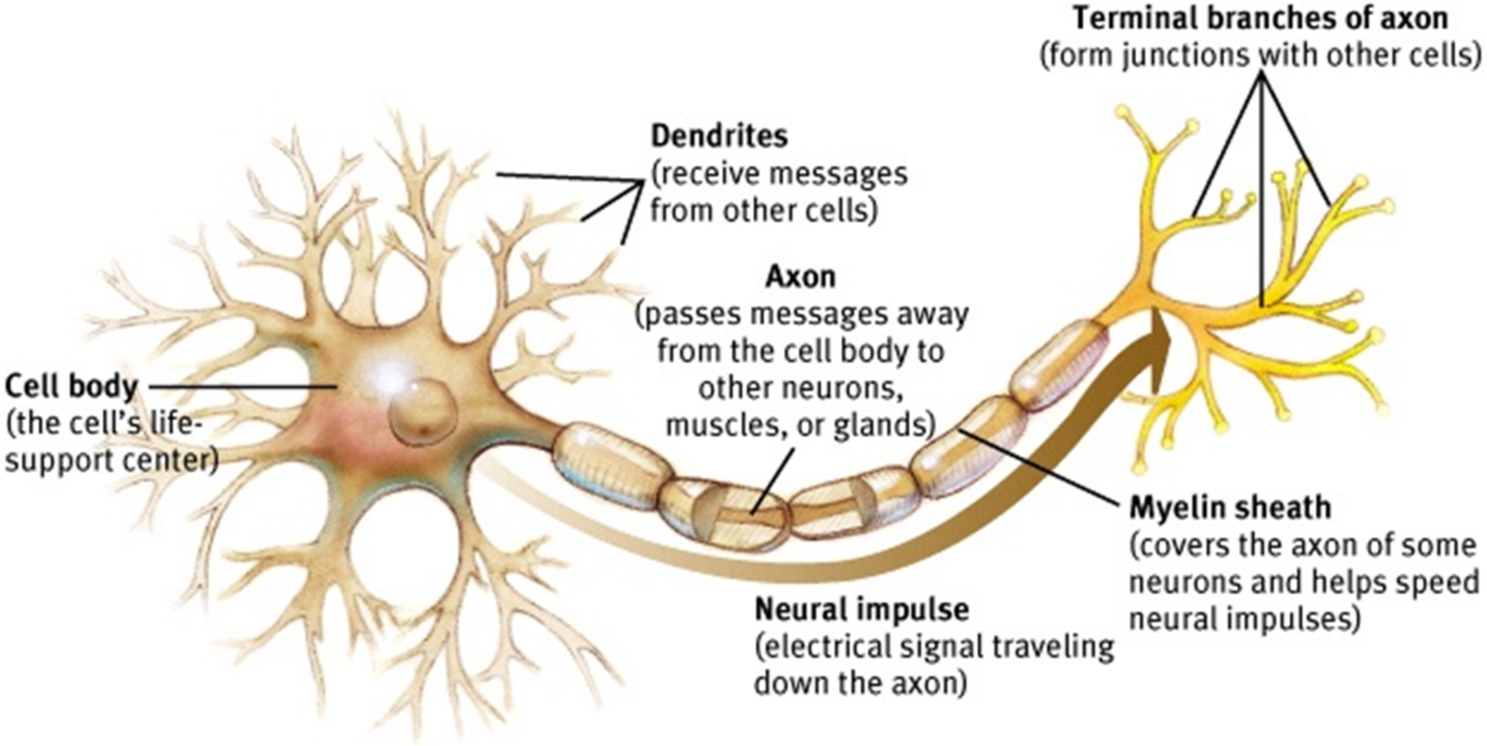
Yann LeCun

Geoff Hinton

Yoshua Bengio

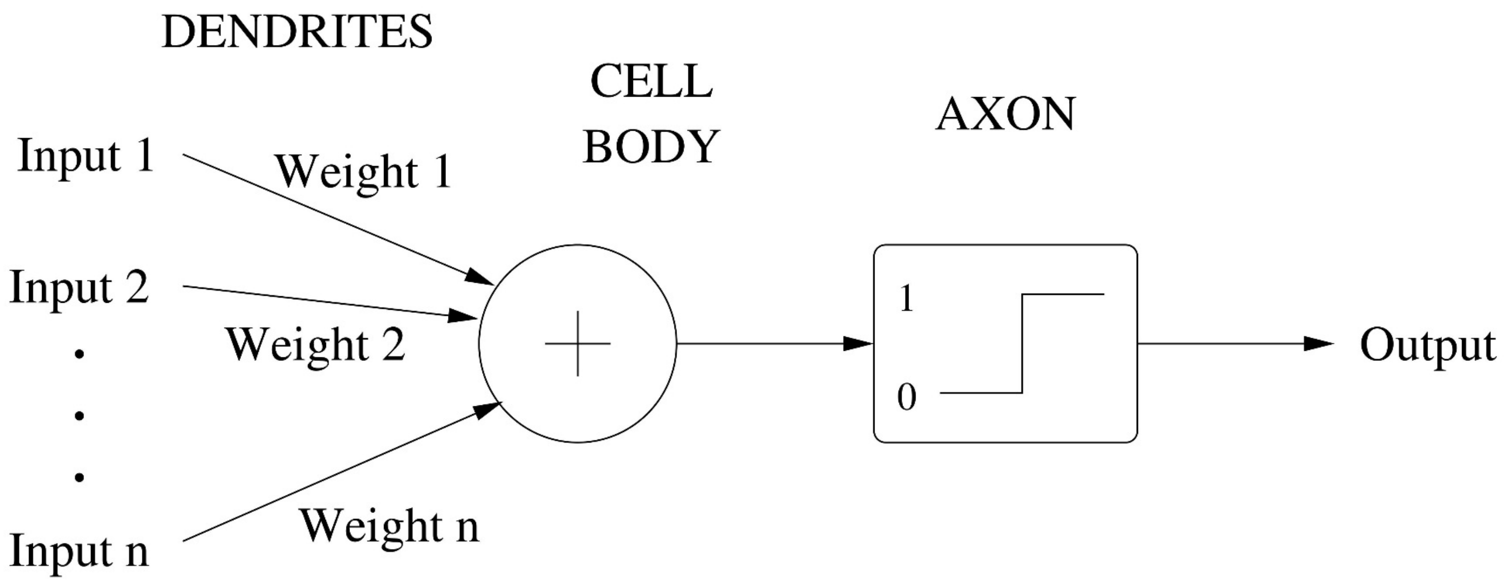
Source: P. Domingos

A Neuron



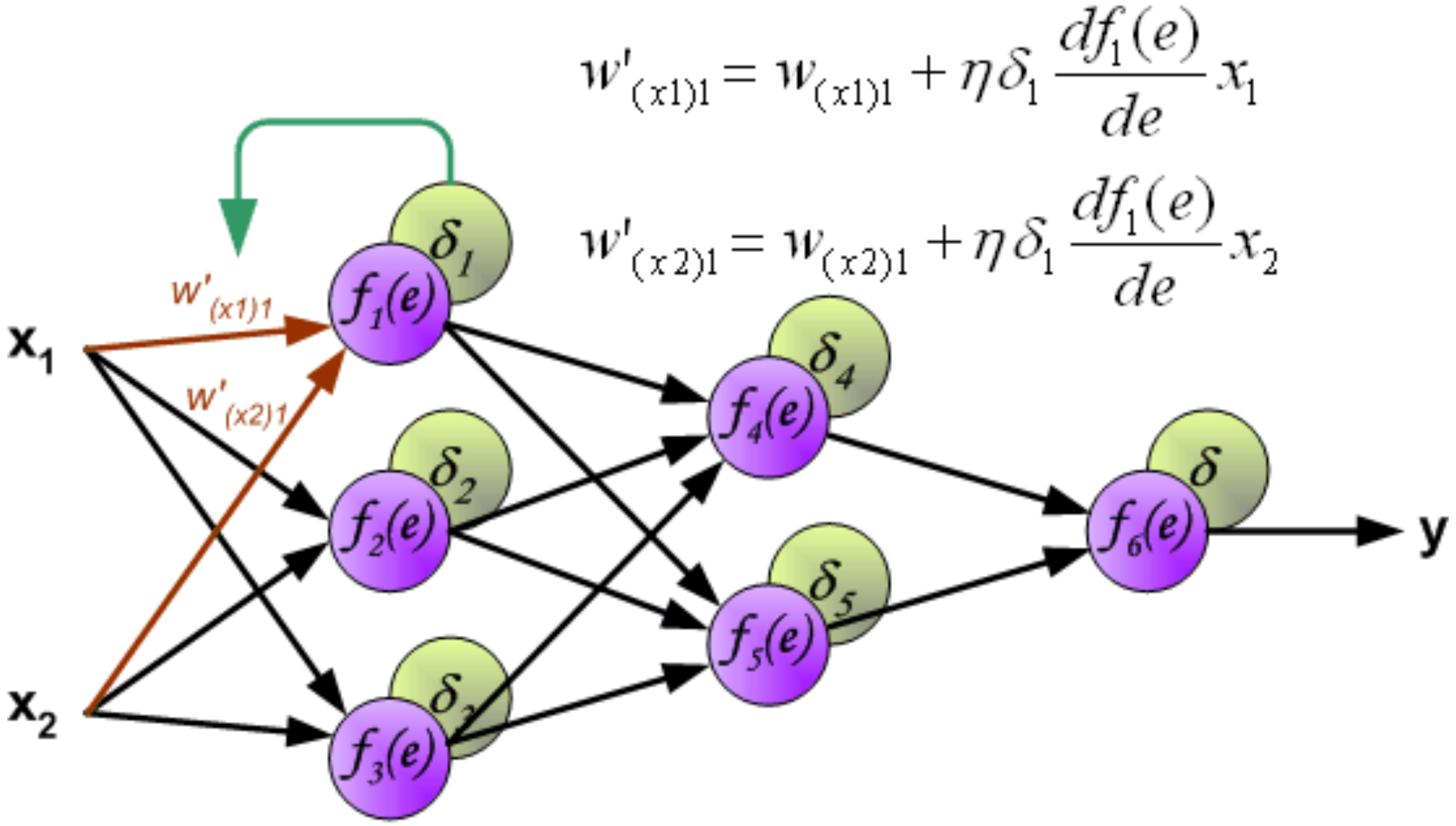
Source: P. Domingos

An Artificial Neuron



Source: P. Domingos

Backpropagation



Source: P. Domingos

Evolutionaries

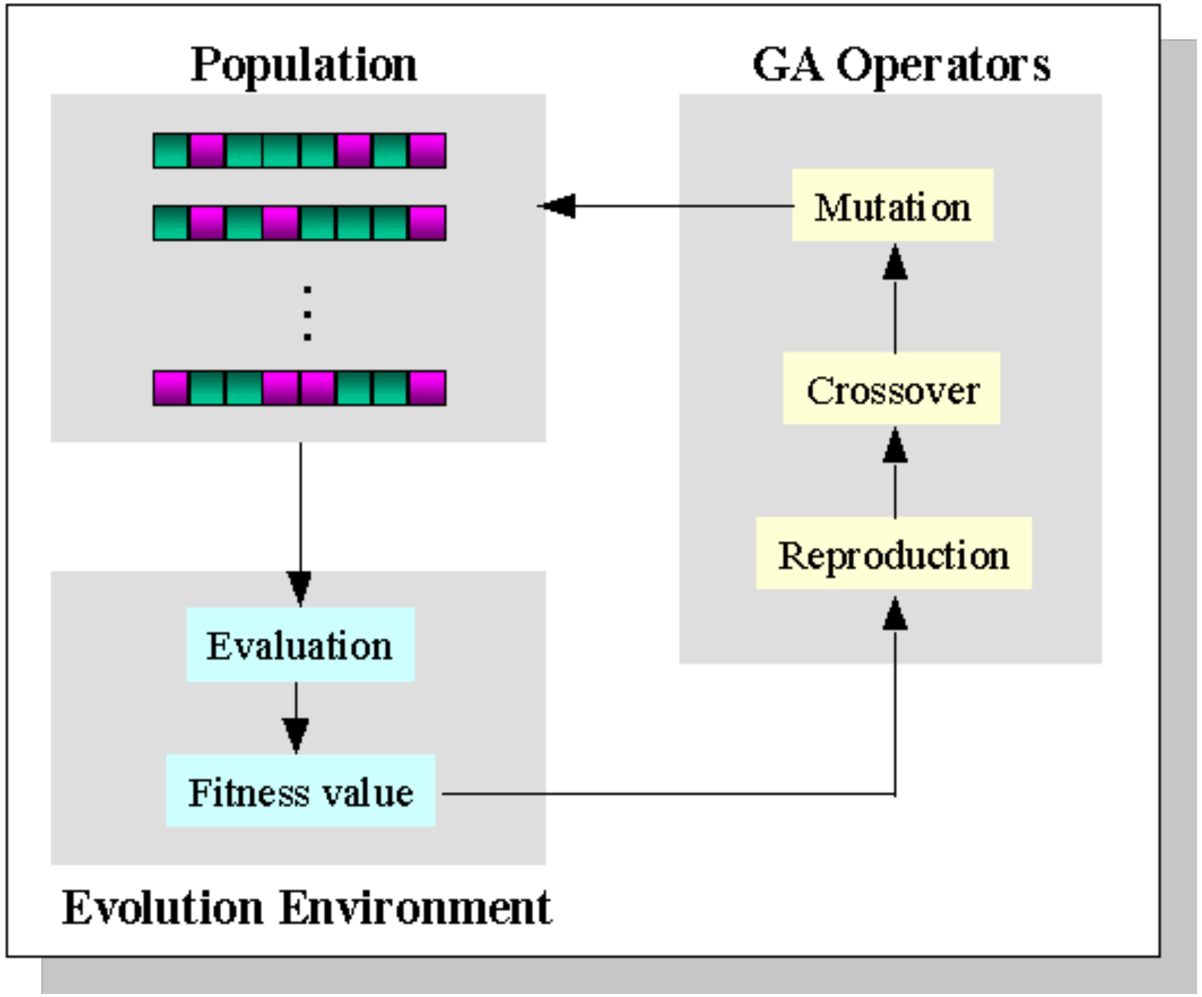
John Koza

John Holland

Hod Lipson

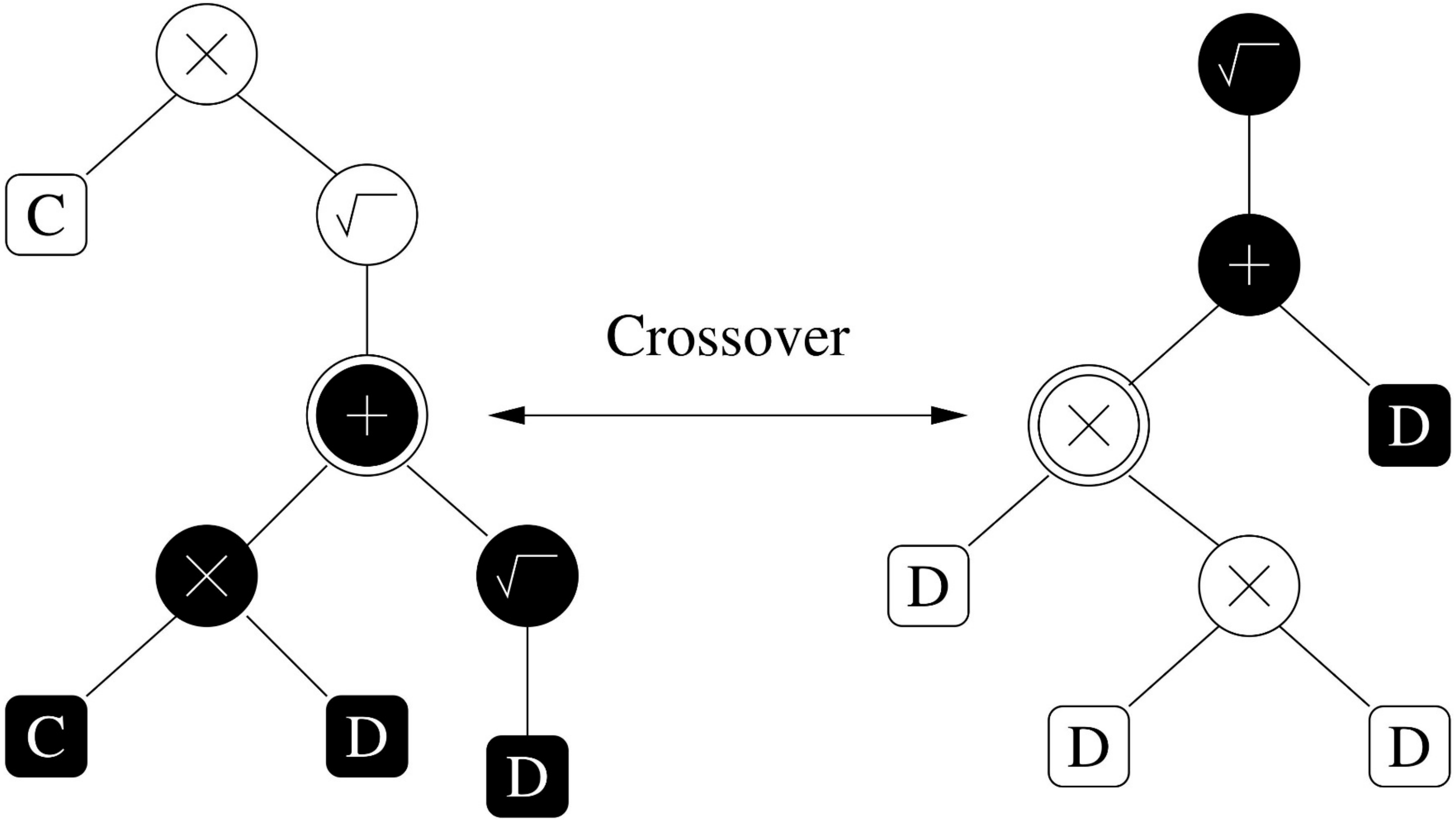
Source: P. Domingos

Genetic Algorithms



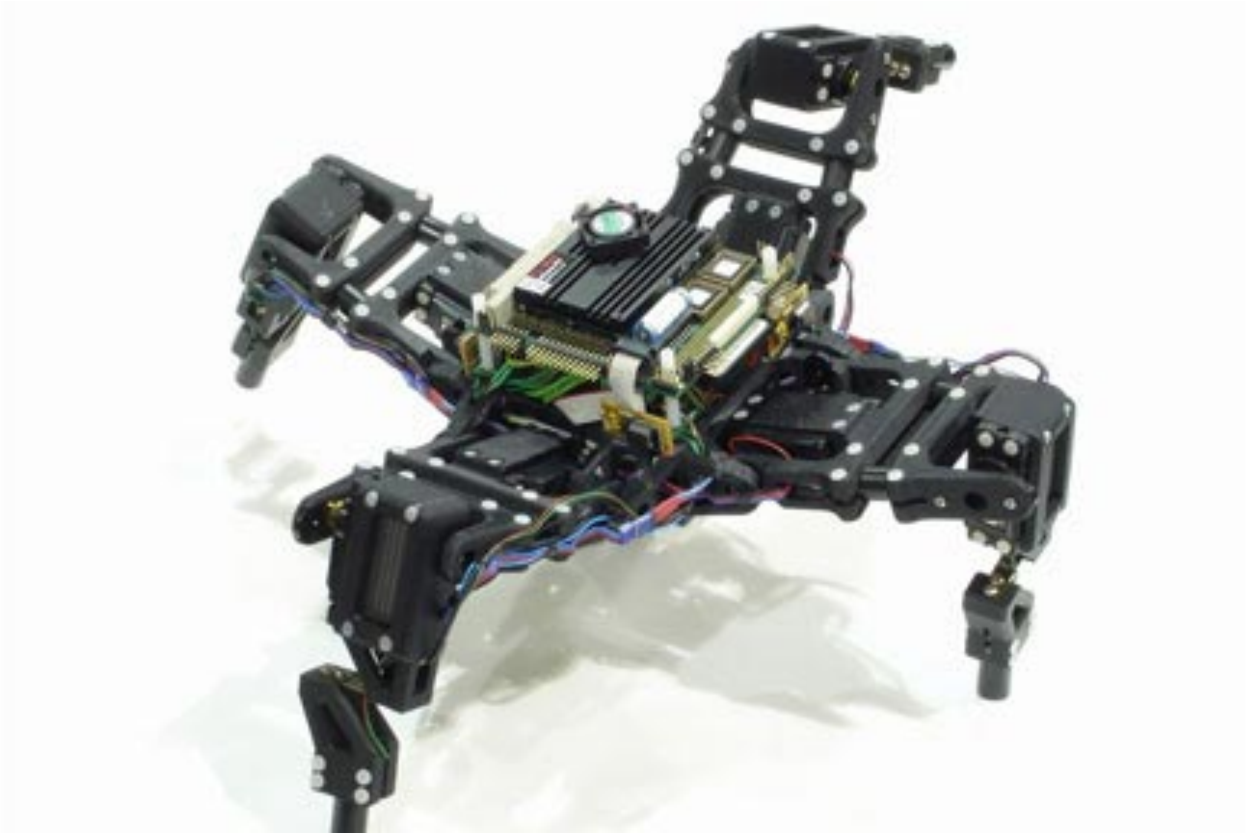
Source: P. Domingos

Genetic Programming



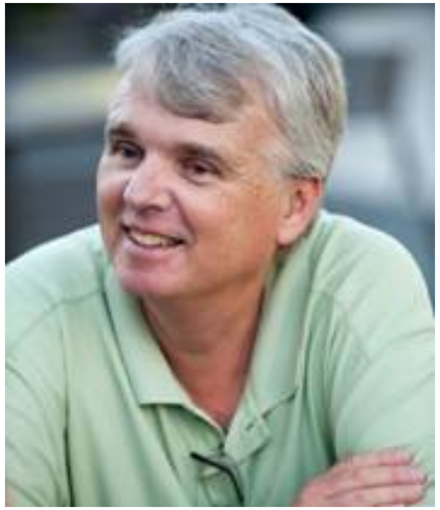
Source: P. Domingos

Evolving Robots



Source: P. Domingos

Bayesians



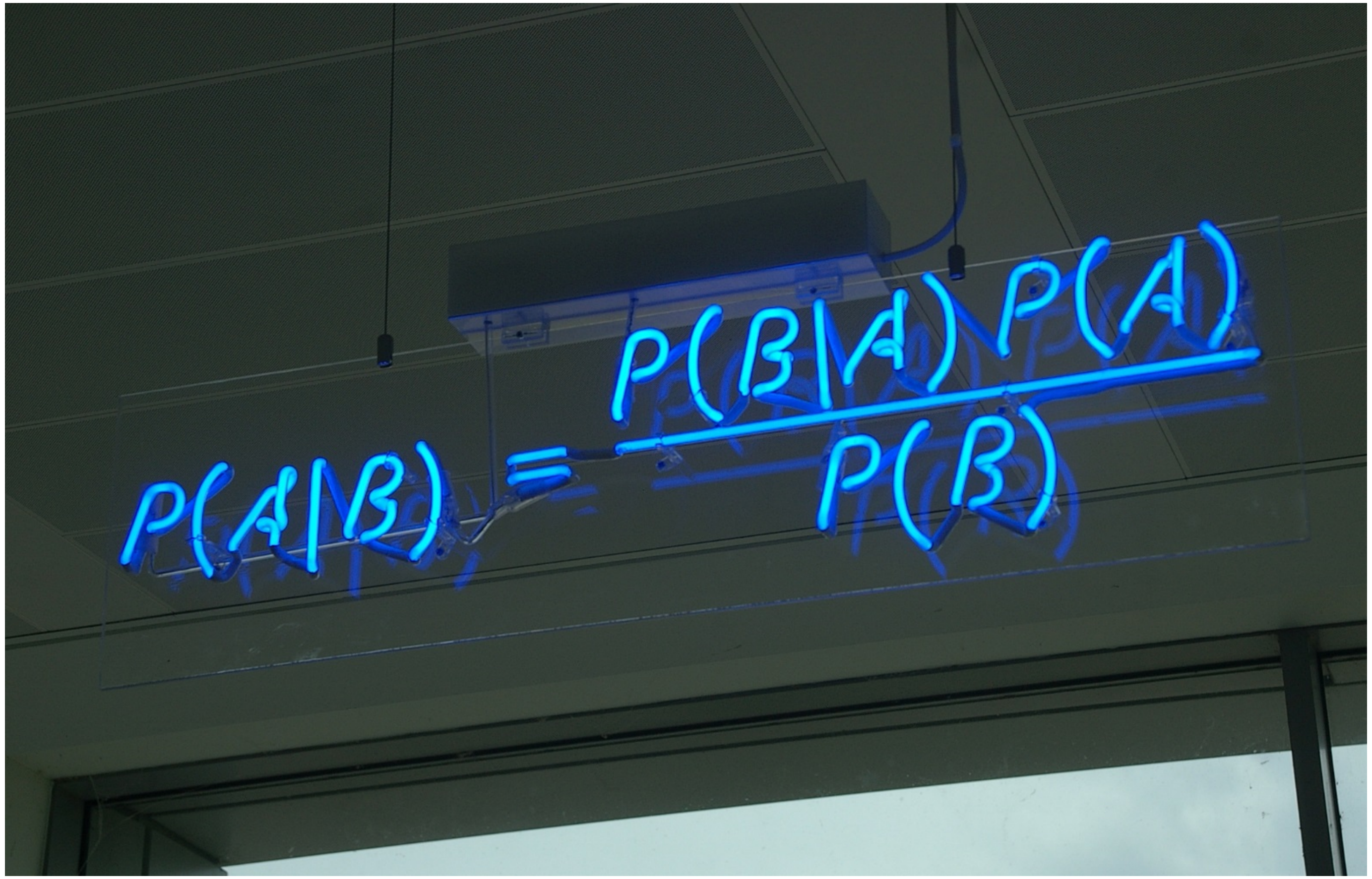
David Heckerman

Judea Pearl

Michael Jordan

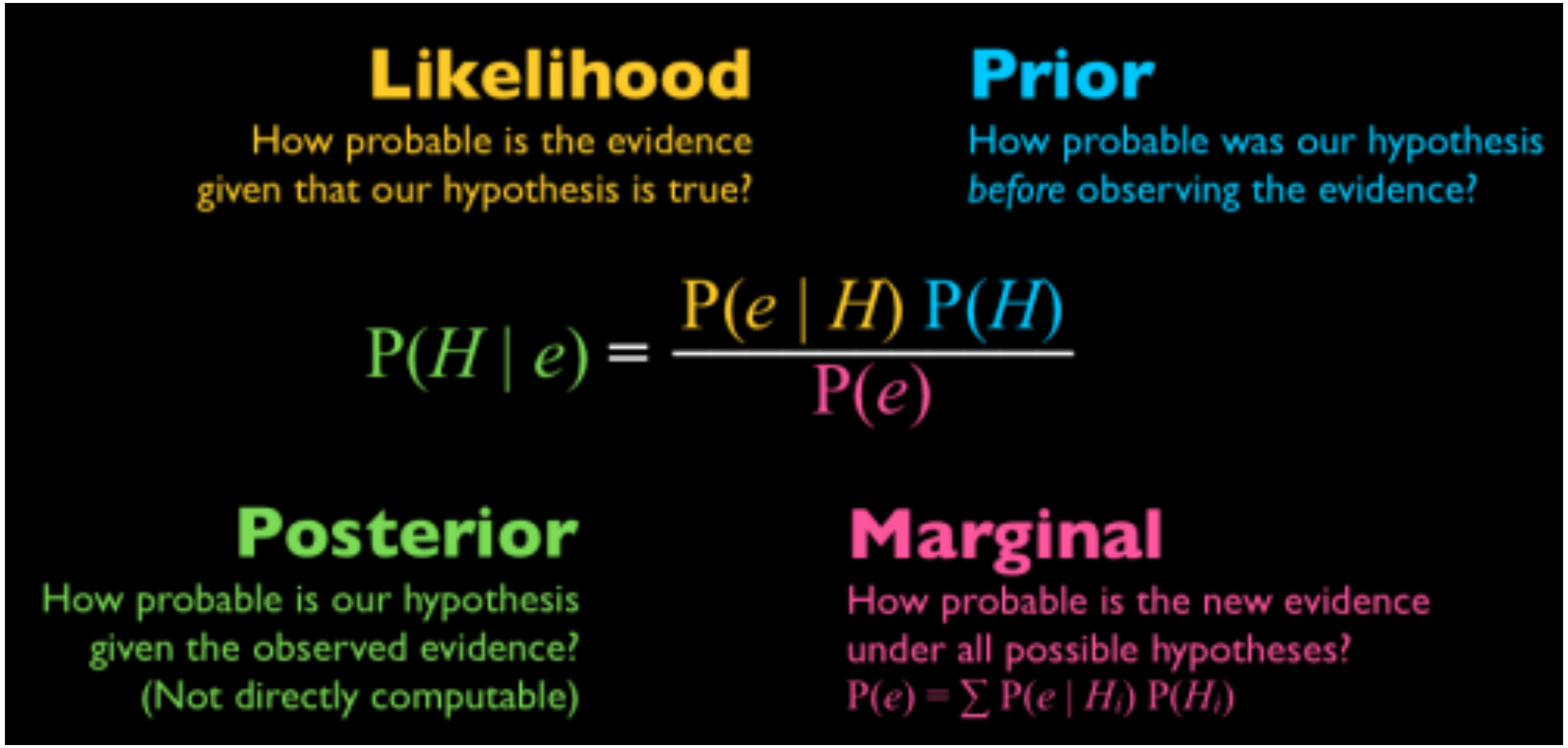
Source: P. Domingos

Probabilistic Inference

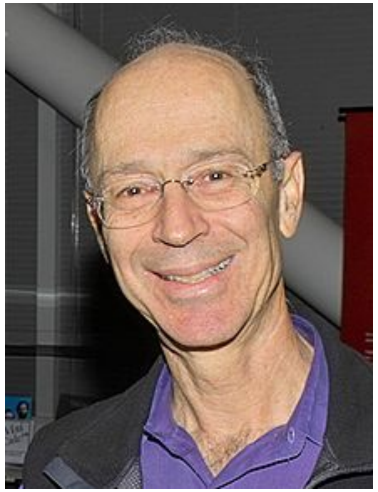


Source: P. Domingos

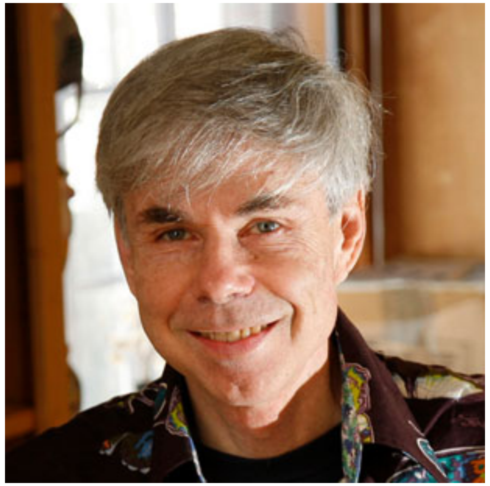
Probabilistic Inference



Analogizers

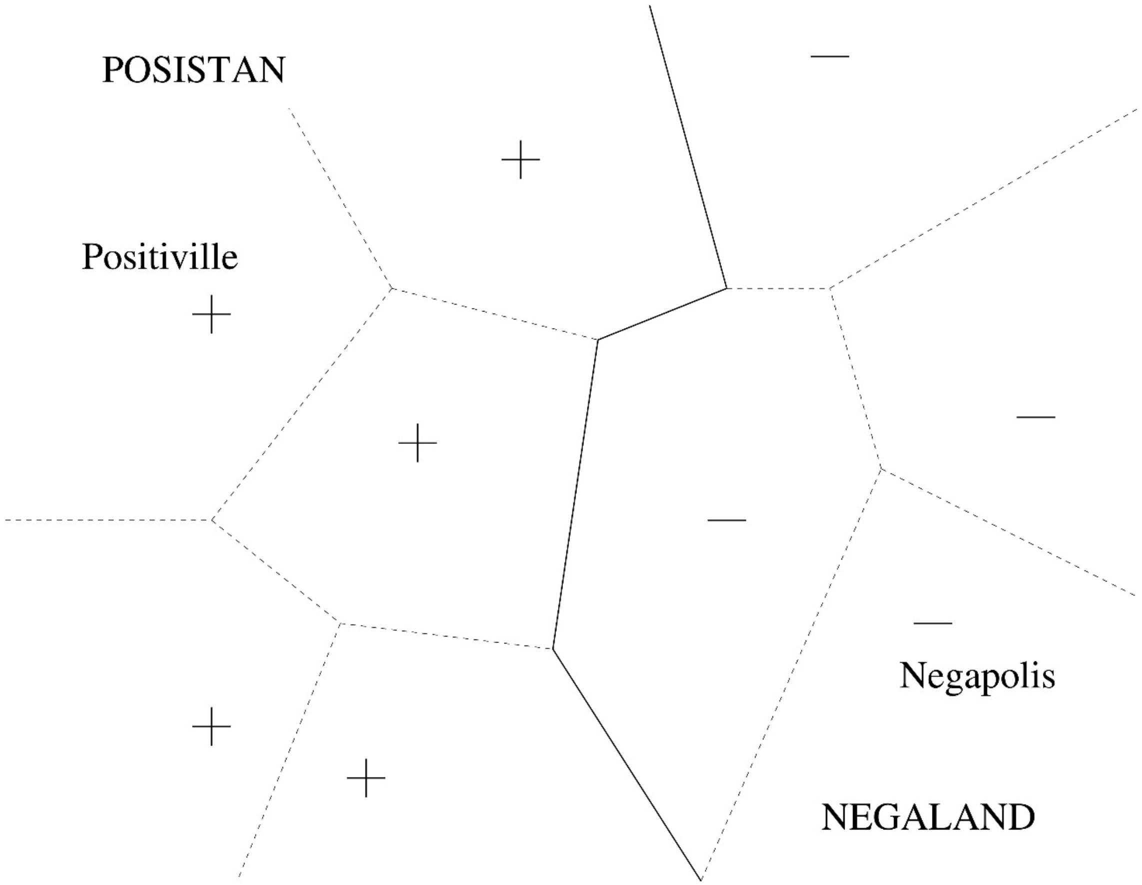

Peter Hart


Vladimir Vapnik


Douglas Hofstadter

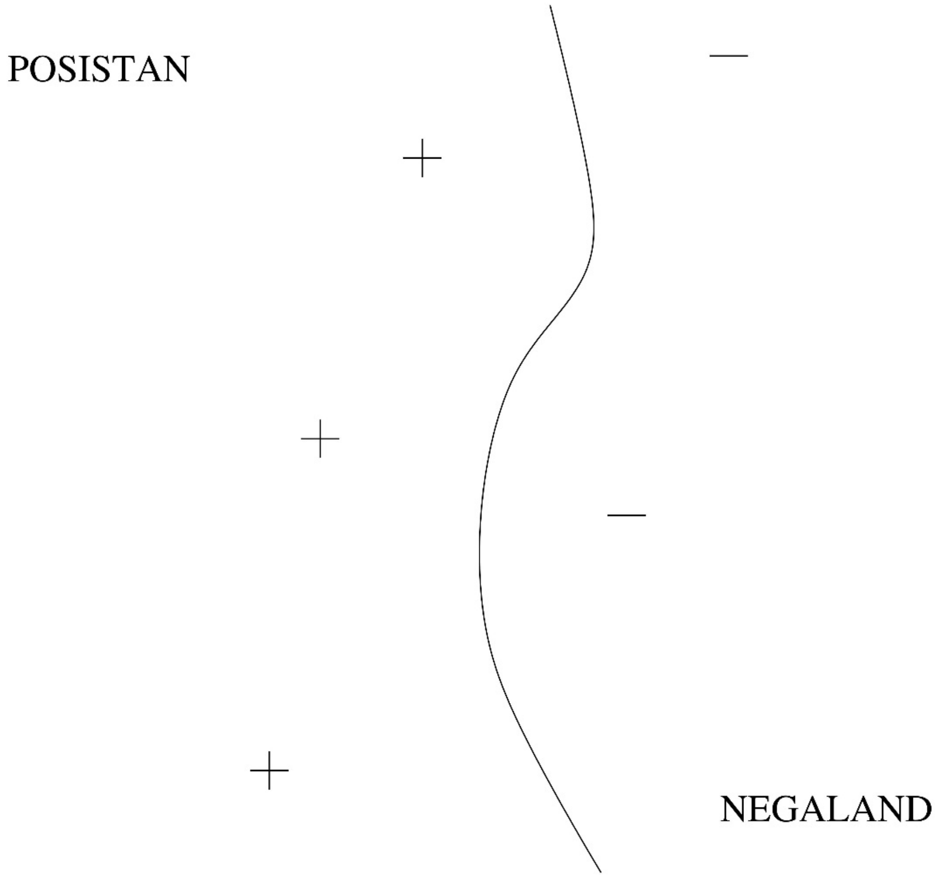
Source: P. Domingos

Nearest Neighbor

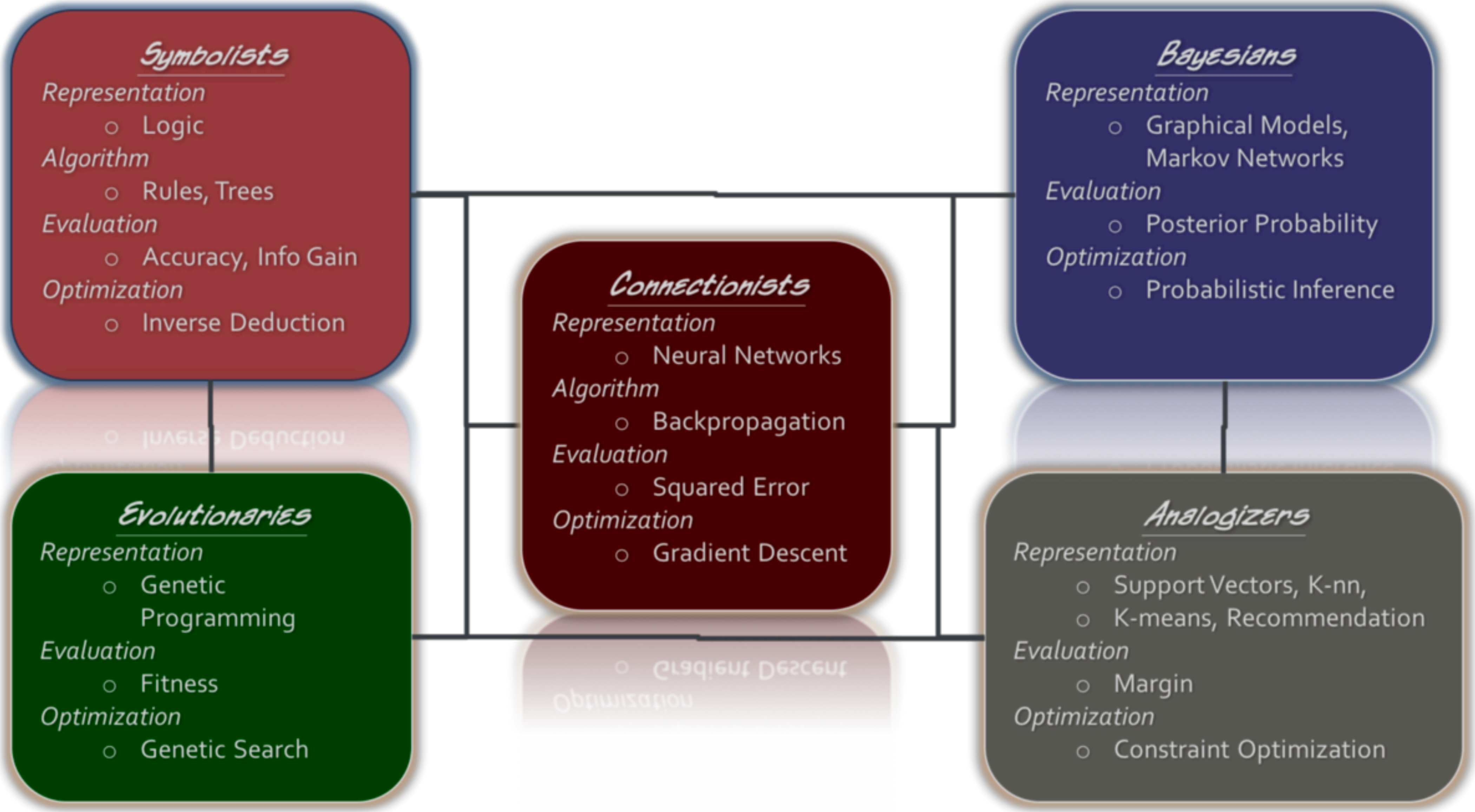


Source: P. Domingos

Kernel Machines



Source: P. Domingos


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

Source: P. Domingos

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

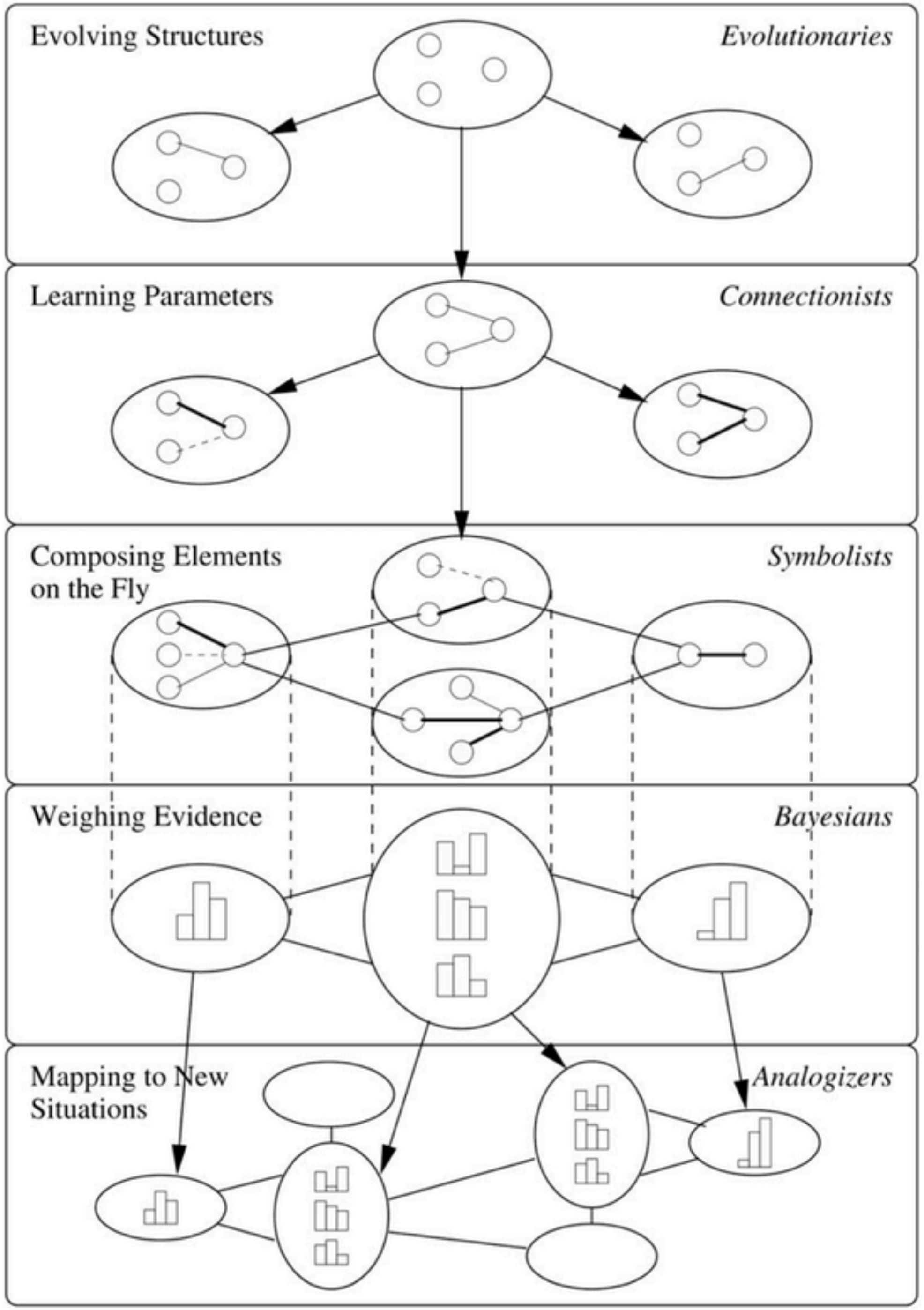
But what we really need is
a single algorithm that solves all five!

Source: P. Domingos

Putting the Pieces Together

- Representation
  - Probabilistic logic (e.g., Markov logic networks)
  - Weighted formulas → Distribution over states
- Evaluation
  - Posterior probability
  - User-defined objective function
- Optimization
  - Formula discovery: Genetic programming
  - Weight learning: Backpropagation

Source: P. Domingos



Source: P. Domingos