CSC520 - Artificial Intelligence

Lecture 1

Dr. Scott N. Gerard

North Carolina State University

Jan 7, 2025

Agenda

- Introductions
- Overview
- History
- State of the Art
- Benefits and Risks
- Logistics

Introductions

- Instructor
  - Dr. Scott N. Gerard - sngerard@ncsu.edu
    - Many decades at IBM
    - Lecturer, North Carolina State University
    - Chief AI Officer at Knowledge Reactor
- TAs
  - Fardin Saad - fsaad@ncsu.edu

Overview

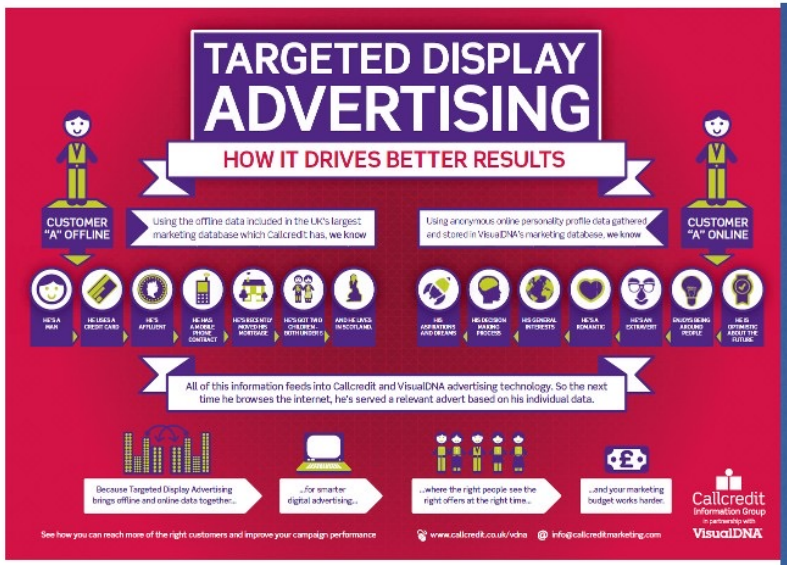
Overview

- Asimov's Three Laws of Robotics
  1. A robot may not injure a human being or, through inaction, allow a human being to come to harm.
  2. A robot must obey orders given it by human beings except where such orders would conflict with the First Law.
  3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

Overview


AltaVista USA
altavista
Web
FIND
Advanced Search
Settings
SEARCH: Worldwide USA RESULTS IN: All languages English, Spanish
Babel Fish Translation
Toolbar Answers Directory Shopping Travel Yellow Pages People Finder More
New Privacy Policy - Please Read
Business Services Submit a Site About AltaVista Privacy Policy Help
© 2010 Overture Services, Inc.


nest


TARGETED DISPLAY
ADVERTISING
HOW IT DRIVES BETTER RESULTS
CUSTOMER "A" OFFLINE
Using the offline data included in the UK's largest marketing database which Callcredit has, we know
Using anonymous online personality profile data gathered and stored in VisualDNA's marketing database, we know
CUSTOMER "A" ONLINE
All of this information feeds into Callcredit and VisualDNA advertising technology. So the next time he browses the internet, he's served a relevant advert based on his individual data.
Because Targeted Display Advertising brings offline and online data together...
...for smarter digital advertising...
...where the right people see the right offers at the right time...
...and your marketing budget works harder.
Callcredit
Information Group
VisualDNA
See how you can reach more of the right customers and improve your campaign performance
www.callcredit.co.uk/vdna
info@callcreditmarketing.com

Overview

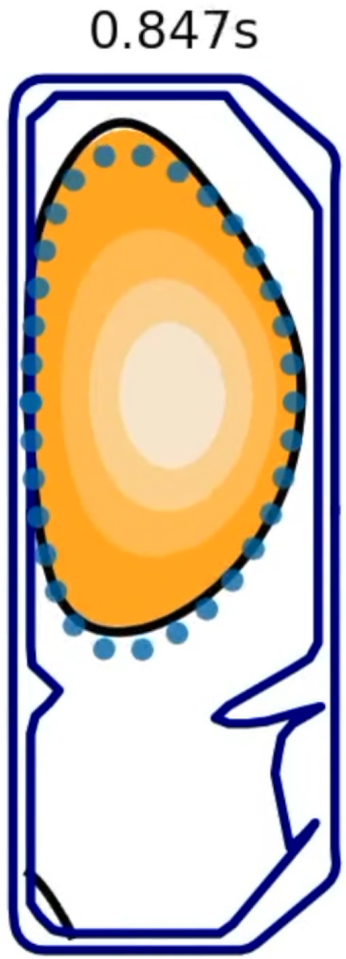
Article

Magnetic control of tokamak plasmas through deep reinforcement learning

https://doi.org/10.1038/s41586-021-04301-9
Received: 14 July 2021
Accepted: 1 December 2021
Published online: 16 February 2022
Open access
Check for updates

Jonas Degrave[1,3], Federico Felici[2,3], Jonas Buchli[1,3], Michael Neunert[1,3], Brendan Tracey[1,3], Francesco Carpanese[1,2,3], Timo Ewalds[1,3], Roland Hafner[1,3], Abbas Abdolmaleki[1], Diego de las Casas[1], Craig Donner[1], Leslie Fritz[1], Cristian Galperti[2], Andrea Huber[1], James Keeling[1], Maria Tsimpoukelli[1], Jackie Kay[1], Antoine Merle[2], Jean-Marc Moret[2], Seb Noury[1], Federico Pesamosca[2], David Pfau[1], Olivier Sauter[2], Cristian Sommariva[2], Stefano Coda[2], Basil Duval[2], Ambrogio Fasoli[2], Pushmeet Kohli[1], Koray Kavukcuoglu[1], Demis Hassabis[1] & Martin Riedmiller[1,3]

Nuclear fusion using magnetic confinement, in particular in the tokamak configuration, is a promising path towards sustainable energy. A core challenge is to shape and maintain a high-temperature plasma within the tokamak vessel. This requires high-dimensional, high-frequency, closed-loop control using magnetic actuator coils, further complicated by the diverse requirements across a wide range of plasma configurations. In this work, we introduce a previously undescribed architecture for tokamak magnetic controller design that autonomously learns to command the full set of control coils. This architecture meets control objectives specified at a high level, at the same time satisfying physical and operational constraints. This approach has unprecedented flexibility and generality in problem specification and yields a notable reduction in design effort to produce new plasma configurations. We successfully produce and control a diverse set of plasma configurations on the Tokamak à Configuration Variable[1,2], including elongated, conventional shapes, as well as advanced configurations, such as negative



Overview

What is intelligence?

Overview

AI Definitions

- Various definitions of AI can be categorized based on two dimensions
  - Comparison with *humans* versus notion of *rationality*

Overview

AI Definitions

- Various definitions of AI can be categorized based on two dimensions
  - Comparison with *humans* versus notion of *rationality*
  - *Internal* thought processes versus *external* behavior

- Various definitions of AI can be categorized based on two dimensions
  - Comparison with *humans* versus notion of *rationality*
  - *Internal* thought processes versus *external* behavior

Think like humans	**Think rationally**
Cognitive science, "machines with minds"	Logical reasoning, "study of mental faculties through the use of computational models"
Act like humans	**Act rationally**
Turing test, "Machines that perform functions that require intelligence when performed by people"	Agents that do the right thing, "intelligent behavior in artifacts"

Overview

Foundations of AI

Philosophy	logic, reasoning, rationality
Mathematics	formal logic, algorithms, complexity analysis, probability
Economics	decision theory, operations research, utility
Neuroscience	human brain, nervous system
Psychology	knowledge-based agent, human-computer interaction
Computer engineering	efficient computers
Linguistics	natural language processing, knowledge representation

History of AI

- Inception of AI (1943-1956)
  - McCulloch and Pitts: Boolean circuit model of brain
  - Turing test

History of AI

- Inception of AI (1943-1956)
  - McCulloch and Pitts: Boolean circuit model of brain
  - Turing test
- Great expectation (1952-1969)
  - AI programs for games, puzzles, etc. based on microworlds
  - Newell and Simon: General problem solver
  - John McCarthy: Lisp programming language
  - Frank Rosenblatt: Perceptrons, perceptron convergence theorem

History of AI

- Inception of AI (1943-1956)
  - McCulloch and Pitts: Boolean circuit model of brain
  - Turing test
- Great expectation (1952-1969)
  - AI programs for games, puzzles, etc. based on microworlds
  - Newell and Simon: General problem solver
  - John McCarthy: Lisp programming language
  - Frank Rosenblatt: Perceptrons, perceptron convergence theorem
- Dose of reality (1966-1986)
  - Early AI systems failed to scale to the size of real-world problems
  - Knowledge-based expert systems industry boom followed by bust
  - AI Winter

History of AI (Cont.)

- Return of neural networks and statistical approaches (1990-2012)
  - Back-propagation learning algorithm re-invented
  - Probability and statistical methods to address uncertainity
  - Hidden markov models, Bayesian networks, reinforcement learning

History of AI (Cont.)

- Return of neural networks and statistical approaches (1990-2012)
  - Back-propagation learning algorithm re-invented
  - Probability and statistical methods to address uncertainity
  - Hidden markov models, Bayesian networks, reinforcement learning
- Recent history (2012-present)
  - Big data, big compute
  - Deep learning
  - AI used extensively
  - Generative AI (GenAI) and Large Language Models (LLM)

State of the Art

Which of the following can be done at present?

- Play and defeat humans in chess

State of the Art

Which of the following can be done at present?

- Play and defeat humans in chess ✓
- Play and defeat humans in Go

State of the Art

Which of the following can be done at present?

- Play and defeat humans in chess ✓
- Play and defeat humans in Go ✓
- Drive a car autonomously

State of the Art

Which of the following can be done at present?

- Play and defeat humans in chess ✓
- Play and defeat humans in Go ✓
- Drive a car autonomously ✓
- Discover and prove a new mathematical theorem

State of the Art

Which of the following can be done at present?

- Play and defeat humans in chess ✓
- Play and defeat humans in Go ✓
- Drive a car autonomously ✓
- Discover and prove a new mathematical theorem ??
- Design and execute a research program in molecular biology

State of the Art

Which of the following can be done at present?

- Play and defeat humans in chess ✓
- Play and defeat humans in Go ✓
- Drive a car autonomously ✓
- Discover and prove a new mathematical theorem ??
- Design and execute a research program in molecular biology X
- Write an intentionally funny story

State of the Art

Which of the following can be done at present?

- Play and defeat humans in chess ✓
- Play and defeat humans in Go ✓
- Drive a car autonomously ✓
- Discover and prove a new mathematical theorem ??
- Design and execute a research program in molecular biology X
- Write an intentionally funny story ✓
- Buy groceries in a local grocery store

State of the Art

Which of the following can be done at present?

- Play and defeat humans in chess ✓
- Play and defeat humans in Go ✓
- Drive a car autonomously ✓
- Discover and prove a new mathematical theorem ??
- Design and execute a research program in molecular biology X
- Write an intentionally funny story ✓
- Buy groceries in a local grocery store X
- Translate spoken English into spoken Chinese in real time

State of the Art

Which of the following can be done at present?

- Play and defeat humans in chess ✓
- Play and defeat humans in Go ✓
- Drive a car autonomously ✓
- Discover and prove a new mathematical theorem ??
- Design and execute a research program in molecular biology ✗
- Write an intentionally funny story ✓
- Buy groceries in a local grocery store ✗
- Translate spoken English into spoken Chinese in real time ✓
- Perform a complex surgical operation

State of the Art

Which of the following can be done at present?

- Play and defeat humans in chess ✓
- Play and defeat humans in Go ✓
- Drive a car autonomously ✓
- Discover and prove a new mathematical theorem ??
- Design and execute a research program in molecular biology ✗
- Write an intentionally funny story ✓
- Buy groceries in a local grocery store ✗
- Translate spoken English into spoken Chinese in real time ✓
- Perform a complex surgical operation ✗
- Unload a dishwasher and put dishes in right places

State of the Art

Which of the following can be done at present?

- Play and defeat humans in chess ✓
- Play and defeat humans in Go ✓
- Drive a car autonomously ✓
- Discover and prove a new mathematical theorem ??
- Design and execute a research program in molecular biology ✗
- Write an intentionally funny story ✓
- Buy groceries in a local grocery store ✗
- Translate spoken English into spoken Chinese in real time ✓
- Perform a complex surgical operation ✗
- Unload a dishwasher and put dishes in right places ✗

State of the Art (Cont.)

- Robotics
  - Self-driving cars
  - Factory automation
  - Soccer

State of the Art (Cont.)

- Robotics
  - Self-driving cars
  - Factory automation
  - Soccer
- Natural Language Processing
  - Speech recognition, text to speech
  - Question answering, language translation, information retrieval
  - General purpose language models, story generation

State of the Art (Cont.)

- Game playing
  - Deep blue defeated world chess champion Garry Kasparov in 1997
  - Alpha Go defeated European champion Fan Hui in 2015

State of the Art (Cont.)

- Game playing
  - Deep blue defeated world chess champion Garry Kasparov in 1997
  - Alpha Go defeated European champion Fan Hui in 2015
- Computer Vision
  - Image classification
  - Object recognition, face recognition
  - Image segmentation
  - Generative adversarial networks

Benefits and Risks of AI

- Benefits
  - Decrease dangerous and repetitive work
  - Increase production of goods and services
  - Accelerate scientific research: cure diseases, climate change solutions

Benefits and Risks of AI

- Benefits
  - Decrease dangerous and repetitive work
  - Increase production of goods and services
  - Accelerate scientific research: cure diseases, climate change solutions
- Risks
  - Lethal autonomous weapons can eliminate human targets
  - Mass surveillance using phone, video, emails, etc.
  - Biased decision making
  - Impact on employment
  - Use of Copyrighted/Intellectual Property materials for ML training
  - Safety critical applications: self-driving cars, water supply mgmt., etc.
  - Hackers using AI for cybersecurity attacks

Benefits and Risks of AI

- Benefits
  - Decrease dangerous and repetitive work
  - Increase production of goods and services
  - Accelerate scientific research: cure diseases, climate change solutions
- Risks
  - Lethal autonomous weapons can eliminate human targets
  - Mass surveillance using phone, video, emails, etc.
  - Biased decision making
  - Impact on employment
  - Use of Copyrighted/Intellectual Property materials for ML training
  - Safety critical applications: self-driving cars, water supply mgmt., etc.
  - Hackers using AI for cybersecurity attacks

Benefits and Risks of AI

- Trustworthy AI
  - Correctness
  - Privacy
  - Bias
  - Explainability
  - Ethics
  - Governance

Course

What will we cover in this course?

- Fundamentals of AI
  - Build agents that act rationally i.e. maximize expected utility

Course

What will we cover in this course?

- Fundamentals of AI
  - Build agents that act rationally i.e. maximize expected utility
- Agents, problem solving, search, knowledge, planning, ML basics

Course

What will we cover in this course?

- Fundamentals of AI
  - Build agents that act rationally i.e. maximize expected utility
- Agents, problem solving, search, knowledge, planning, ML basics
- This is not a Machine Learning or Deep Learning course!

Course Topics

- Agents: Problems, problem solving.
- Search: Uninformed, informed, adversarial.
- Optimization: Local search and adaptation.
- Constraints: Solving problems with constraints.
- Knowledge: Logic, knowledge-based agents.
- Planning: Semi-constrained search, scheduling.
- Uncertainty: Probabilistic models, decision-making.
- ML: Machine Learning, deep learning.
- NLP: Basics of NLP in AI.
- CV: Basics of CV in AI.
- Risks and Benefits:

Large Language Models

- Kind of Neural Network
- Attention/Transformer mechanism

Large Language Models

- Kind of Neural Network
- Attention/Transformer mechanism
- Not the only AI technology
  - They hallucinate
  - Poor reasoners

Large Language Models

- Kind of Neural Network
- Attention/Transformer mechanism
- Not the only AI technology
  - They hallucinate
  - Poor reasoners

Agentic AI

- All agents wrap an LLM
- Not the only AI agent technology

Course Topics

Generative AI

Large Language Models

- Kind of Neural Network
- Attention/Transformer mechanism
- Not the only AI technology
  - They hallucinate
  - Poor reasoners

Agentic AI

- All agents wrap an LLM
- Not the only AI agent technology

LLMs are not covered in this course

Machine
Learning
Tribes

Domingos, *The Master Algorithm*

Symbolist
Logic

Machine
Learning
Tribes

Domingos, *The Master Algorithm*

Course Topics

Five Tribes of Machine Learning

Symbolist
Logic

Connectionist
Neural Network

Machine
Learning
Tribes

Domingos, *The Master Algorithm*

Course Topics

Five Tribes of Machine Learning

Symbolist
Logic

Connectionist
Neural Network

Machine
Learning
Tribes

Bayesian
Probability

Domingos, *The Master Algorithm*

Course Topics

Five Tribes of Machine Learning

Symbolist
Logic

Connectionist
Neural Network

Machine
Learning
Tribes

Bayesian
Probability

Evolutionary
Survival of the Fittest

Domingos, *The Master Algorithm*

Course Topics

Five Tribes of Machine Learning

Symbolist
Logic

Connectionist
Neural Network

Machine
Learning
Tribes

Bayesian
Probability

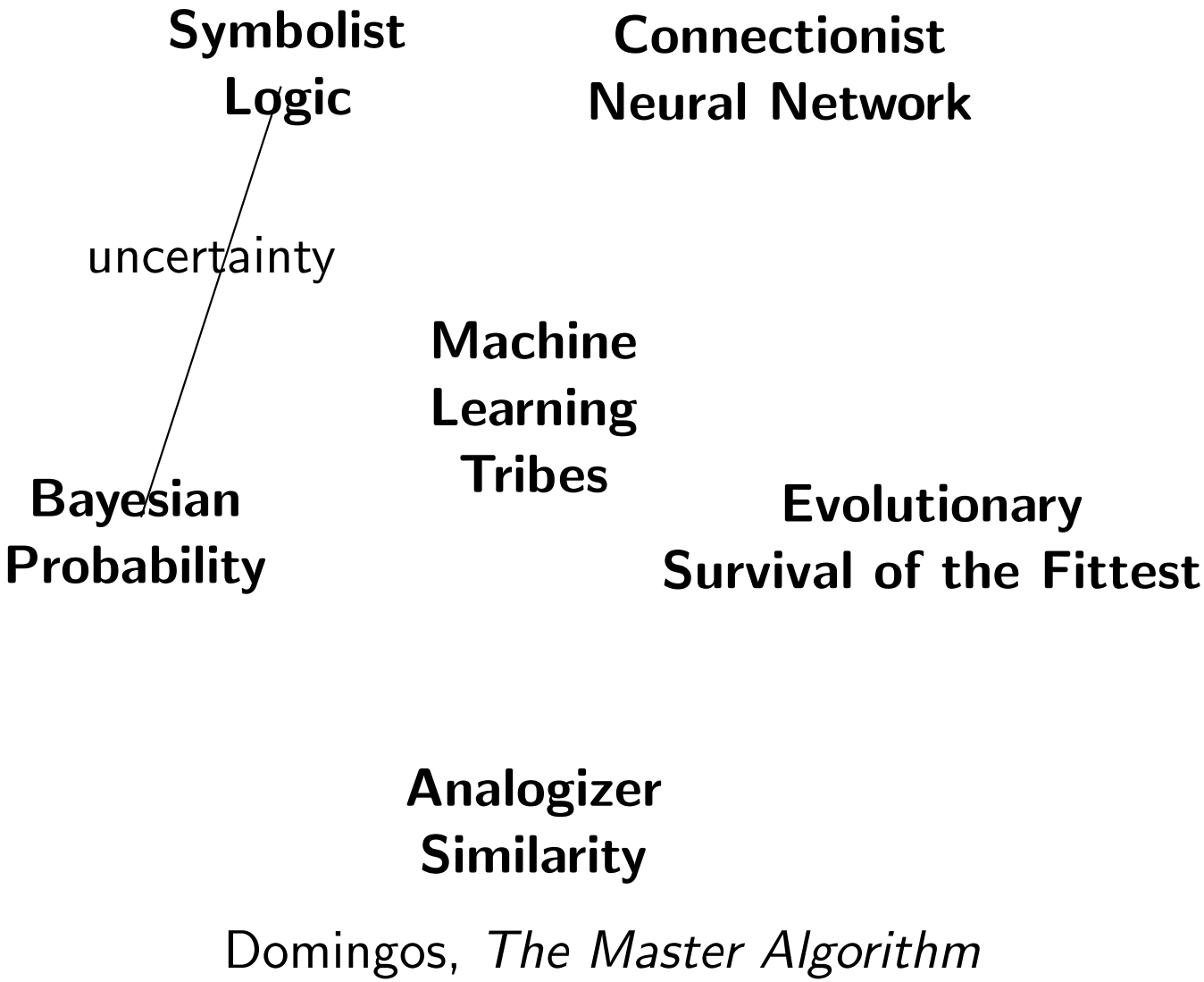
Evolutionary
Survival of the Fittest

Analogizer
Similarity

Domingos, *The Master Algorithm*

Course Topics

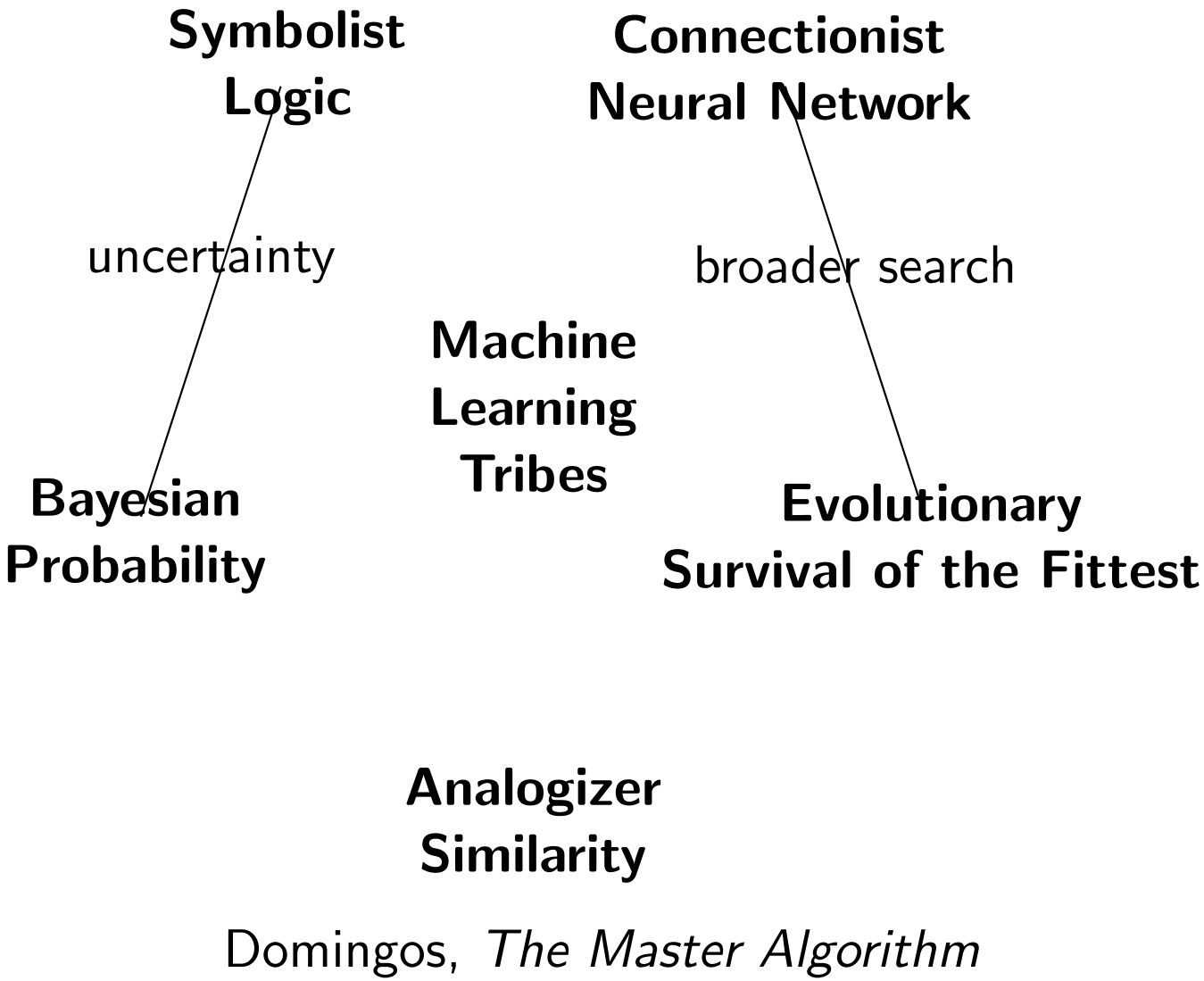
Five Tribes of Machine Learning



Domingos, *The Master Algorithm*

Course Topics

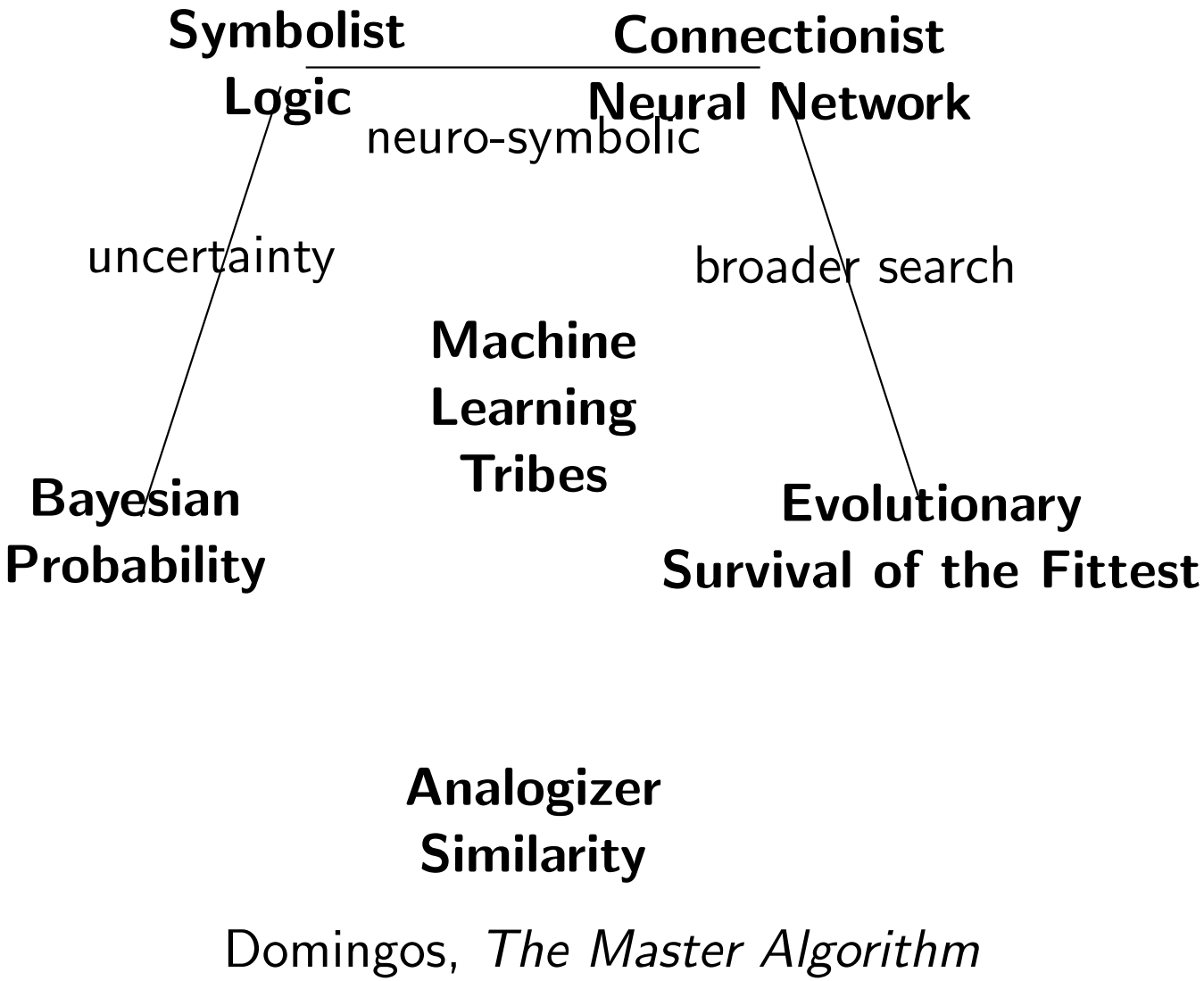
Five Tribes of Machine Learning



Domingos, *The Master Algorithm*

Course Topics

Five Tribes of Machine Learning



Domingos, *The Master Algorithm*

Course Logistics

Logistics