Probabilistic Reasoning and Decision Making *ECE 493 T25*

Mark Crowley

Assistant Professor
Department of Electrical and Computer Engineering
University of Waterloo

mcrowley@uwaterloo.ca
https://markcrowley.ca

@compthink

Outline

- A bit about me and my research
- A brief history of AI as I see it
  - Context for the recent Turing Award
  - What does AI mean now?
  - What is Deep Learning?
  - What is Reinforcement Learning?
- Course Outline
- Questions to you about how this course should proceed
  - Lectures vs discussions
  - Reading
  - Collaborative Notes

About Me

- Mark Crowley (*crohw-lee) <he/him>*
- Academics
  - B.A. : York University Computer Science (1999)
  - PhD : Computer Science at the University of British Columbia (2011)
  - Postdoc : Oregon State University - On Computational Sustainability
- IBM Software Developer (loong ago)
- Assistant Professor (since 2015)
  - University of Waterloo – ECE Department
  - **UW ECE Machine Learning Lab**
    - (https://uwaterloo.ca/scholar/mcrowley/lab)
  - **Waterloo.ai** : Waterloo Artificial Intelligence Institute (http://waterloo.ai)
  - **Element AI**: Faculty Research Fellow

Research Vision

*To augment human **decision making** in complex domains and environments in a **dependable** and **transparent** way.*

Domains (where does the complexity come from?)

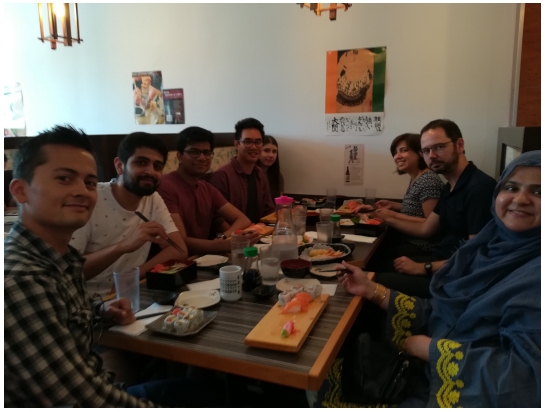
- Spatially Spreading processes in natural systems
- Classification and Anomaly Detection
  - for Streaming Data
  - in medical Imaging
- Automotive Driving Behaviour and Control
  - Driver Behaviour Learning
  - Autonomous Driving
  - Object classification and understanding

Methods (how do we solve it?)

- Reinforcement Learning
- Deep Learning
- Ensemble Methods
- Data/Dimensionality Reduction

My Students

UW ECE Machine Learning Lab



- Laura McCrackin – PhD – Medical imaging
- Maria Samad – MASc –Anomaly Detection
- Benyamin Ghojogh – PhD –Theory, Data Reduction
- Dan Chung – MASc – Autonomous Driving
- **Sriram Subramanian – PhD – Multi-Agent RL**
- **Sami Alperen Akgun – MASc – Robot-human Communication**
- Nouha Chatti – MASc – RL for Chemistry/Material Design

Graduated:

- Sushrut Bhalla – MASc – Combustion, Autonomous Driving → Amazon
- **Pardis Zohouri** – MASc - LRCN for Spatial Spread
- **Sahil Perreira** – MASc – Driver Behaviour Learning → Amazon
- **Juan Carrillo** – MASc – Winter Road Weather Condition Monitoring from Traffic Cameras
- Jaspreet Singh Sambee – MEng – Autonomous Driving → Amazon
- Olivier Nguyen – MASc – Twitter Mining for Health Data → Element AI

Outline

- A bit about me and my research
- A brief history of AI as I see it
  - Context for the recent Turing Award
  - What does AI mean now?
  - What is Deep Learning?
  - What is Reinforcement Learning?
- Course Outline
- Questions to you about how this course should proceed
  - Lectures vs discussions
  - Reading
  - Collaborative Notes

What is Artificial Intelligence?

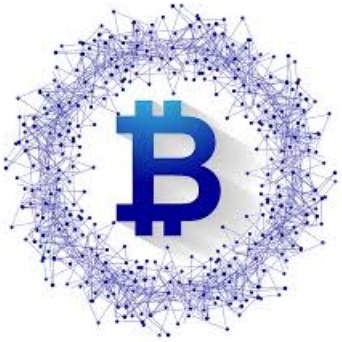
"Artificial Intelligence : any algorithm to enable computers to perform actions we define as requiring intelligence." - Me

- Good Old Fashioned AI
  - Logical rules, planning, search
- Evolutionary computation (e.g. genetic algorithms)
- Reasoning about Probabilities
- Computer Vision
- Natural Language Processing
- Robotics
- Machine Learning
- Decision Making (Reinforcement Learning)

What does Machine Learning mean to you?

- What's the first thing you think of when someone says "Artificial Intelligence" or "Machine Learning"?
- What do you worry about?
- What is exciting about it?

"Banned" Words



The Elephant in the ~~Room~~ Zoom

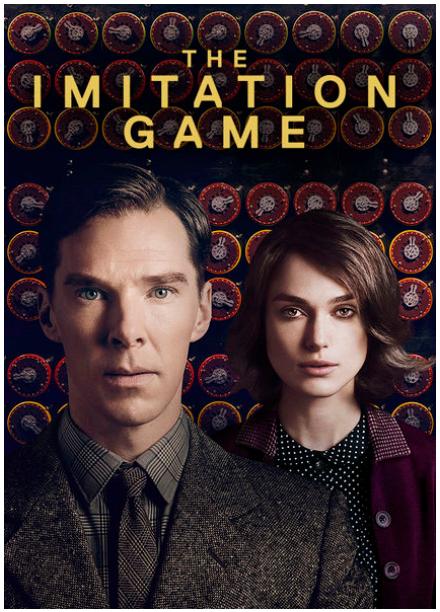
There is one thing which as AI researchers/practitioners you'll always need to be ready to talk about…

A Short History of AI

- 40's – A simple model of the neuron introduced by McCulloh and Pitt as a summing and thresholding device
- 1950 – Alan Turing defines a test for intelligence based on conversation via text interface
- Dartmouth Conference 1956: the birth of AI – Minsky, McCarthy, Shannon, Rochester
- **Period of optimism:** research on logic, search, planning, games
- 1958 – Logistic Regression developed (Cox)
- 1958 – Rosenblatt in 1958 introduced the **Perceptron**, a two layer network (one input layer and one output node with a bias in addition to the input features.
- **1969** – Marvin Minsky: Perceptrons are 'just' linear, AI goes logical, beginning of first "AI Winter"

Alan Turing

Alan M. Turing (1912-1954) was a British mathematician who defined the mathematical foundations of computer science.

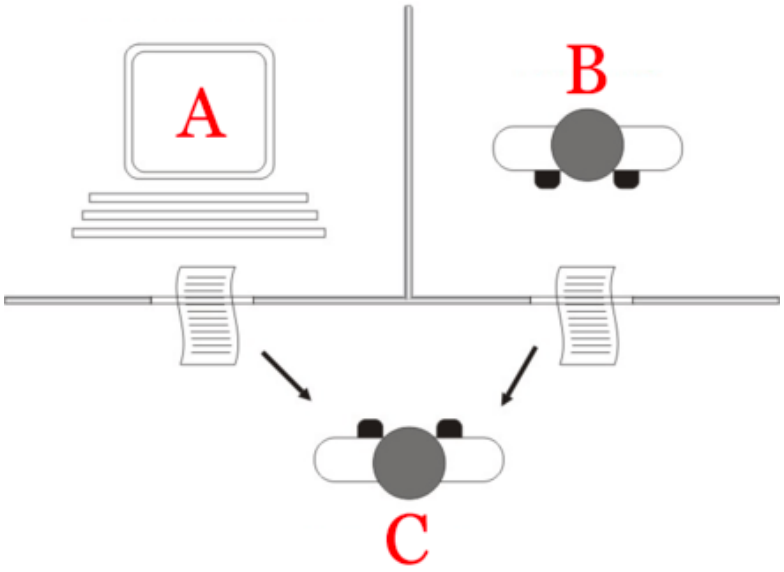


Huge contributions in:

- Algorithm complexity (*how long will it take?*)
- Definition of computing (*what is a computer?)*
- Codebreaking (*what are they saying?*)
- Artificial Intelligence (*what is "thinking"?*)

The Turing Test

- Thought experiment devised by Alan Turing in 1950 to define how to test if a machine passed a *baseline level* of intelligence
- *If a computer can fool you into thinking it is intelligent, maybe it is…*



https://en.wikipedia.org/wiki/Turing_test

Today:

- Chatbot challenges
- CAPTCHA Images

A Short History of AI

- 1980's - *Expert systems and fuzzy logic*
  - Another boom
  - Using logic and human knowledge to build reasoning systems to help solve hard problems
  - No learning or data involved, just complex programming and rules
  - Medical diagnosis, weather, planning
  - **Overpromised again!**
  - **Second AI Winter**
- Rest of 90s:
  - Probabilistic Modelling, Bayesian Networks
  - Kernels can do anything! (they're *really* tricky though)
  - Neural Networks out of favour
  - May 1997 IBM Deep Blue defeated chess grandmaster Garry Kasparov

A Short History of Neural Networks

- 2012 – Google Cat Youtube, speech recognition, self driving cars, computer defeats regional Go champion, ...
- 2014 – GoogLeNet added many layers and introduced inception modules
- 2014 – Generative Adversarial Networks (GANs) introduced.
- 2015 – Microsoft algorithm beats human performance at ImageNet challenge.
- **2016 – AlphaGo defeats one of best world players of Go Lee Sedol using Deep Reinforcement Learning.**
- **2016 – Deep Mind introduces A3C Deep RL algorithm that can learn to play Atari games from images by playing with no instructions.**
- **2018 – Deep Mind Starcraft victories against world-class human players?**

Hot off the Press!

- The “Nobel Prize” of computing
- $1 million sponsored by Google
- Named after **Alan M. Turing**

Yoshua Bengio

Geoffrey E Hinton

Yann LeCun

FATHERS OF THE DEEP LEARNING REVOLUTION RECEIVE ACM A.M. TURING AWARD

Bengio, Hinton, and LeCun Ushered in Major Breakthroughs in Artificial Intelligence

ACM named Yoshua Bengio, Geoffrey Hinton, and Yann LeCun recipients of the 2018 ACM A.M. Turing Award for conceptual and engineering breakthroughs that have made deep neural networks a critical component of computing. Bengio is Professor at the University of Montreal and Scientific Director at Mila, Quebec’s Artificial Intelligence Institute; Hinton is VP and Engineering Fellow of Google, Chief Scientific Adviser of The Vector Institute, and University Professor Emeritus at the University of Toronto; and LeCun is Professor at New York University and VP and Chief AI Scientist at Facebook.

Working independently and together, Hinton, LeCun and Bengio developed conceptual foundations for the field, identified surprising phenomena through experiments, and contributed engineering advances that demonstrated the practical advantages of deep neural networks. In recent years, deep learning methods have been responsible for astonishing breakthroughs in computer vision, speech recognition, natural language processing, and robotics—among other applications.

While the use of artificial neural networks as a tool to help computers recognize patterns and simulate human intelligence had been introduced in the 1980s, by the early 2000s, LeCun, Hinton and Bengio were among a small group who remained committed to this approach. Though their efforts to rekindle the AI community’s interest in neural networks were initially met with skepticism, their ideas recently resulted in major technological advances, and their methodology is now the dominant paradigm in the field.

The “Godfathers” of Deep Learning

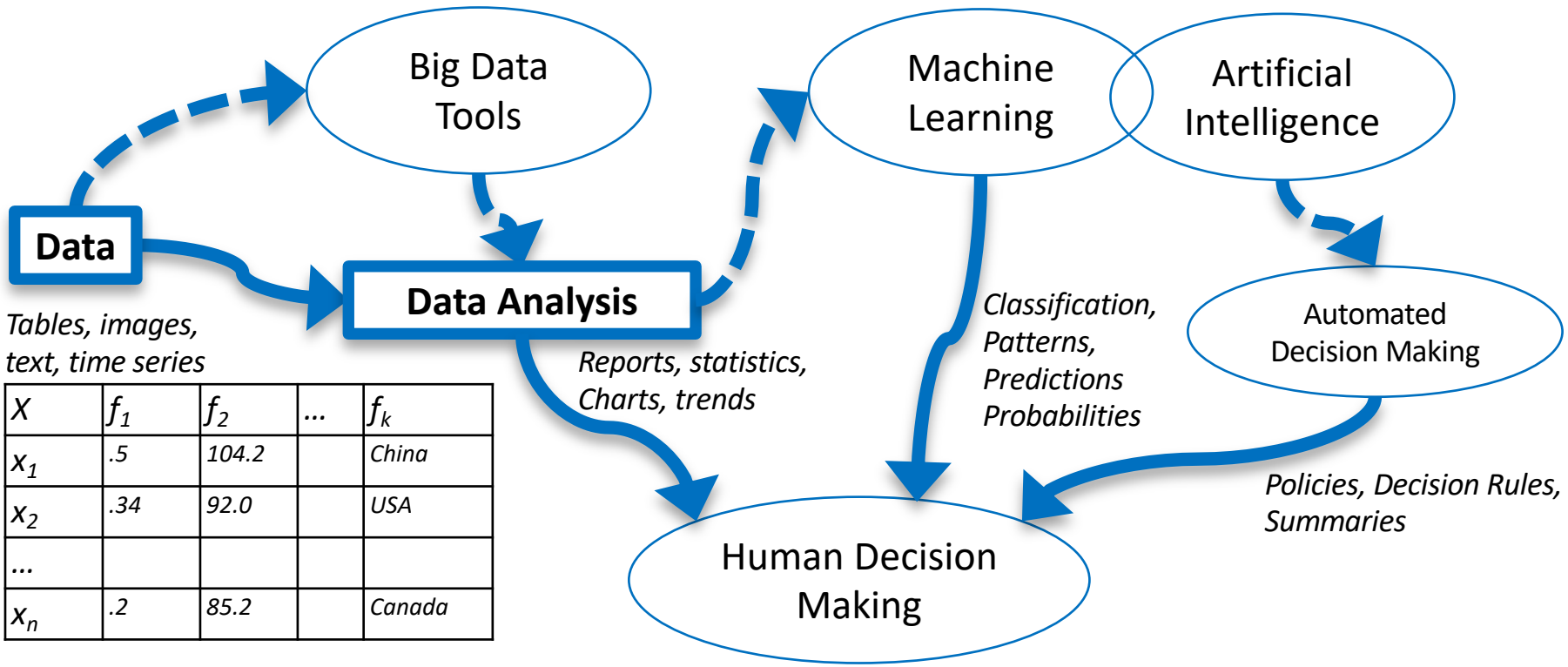
- **Geoffrey Hinton**
  - UofT, now Google Research
  - early work on **back-propagation**
  - Restricted Boltzman Machines - **first deep architecture**, idea for unsupervised pretraining of layers for RBMs was used for first successful use of deep networks.
  - Several applications of Deep Learning to image recognition, speech recognition,...
- **Yoshua Bengio**
  - University of Montreal, MILA, ElememntAI
  - ReLU improves efficiency
  - Unsupervised representation learning of autoencoders
  - **word2vec** - learn general representation of words for classification, translation, etc.
  - **General Adversarial Networks** - an exciting new approach to training
- **Yann LeCun**
  - NYU, now Facebook AI Research
  - Postdoc University of Toronto with Geoffrey Hinton
  - Convolutional Neural Networks – using neural networks on **image understanding**

Outline

- A bit about me and my research
- A brief history of AI as I see it
  - Context for the recent Turing Award
  - What does AI mean now?
  - What is Deep Learning?
  - What is Reinforcement Learning?
- Course Outline
- Questions to you about how this course should proceed
  - Lectures vs discussions
  - Reading
  - Collaborative Notes

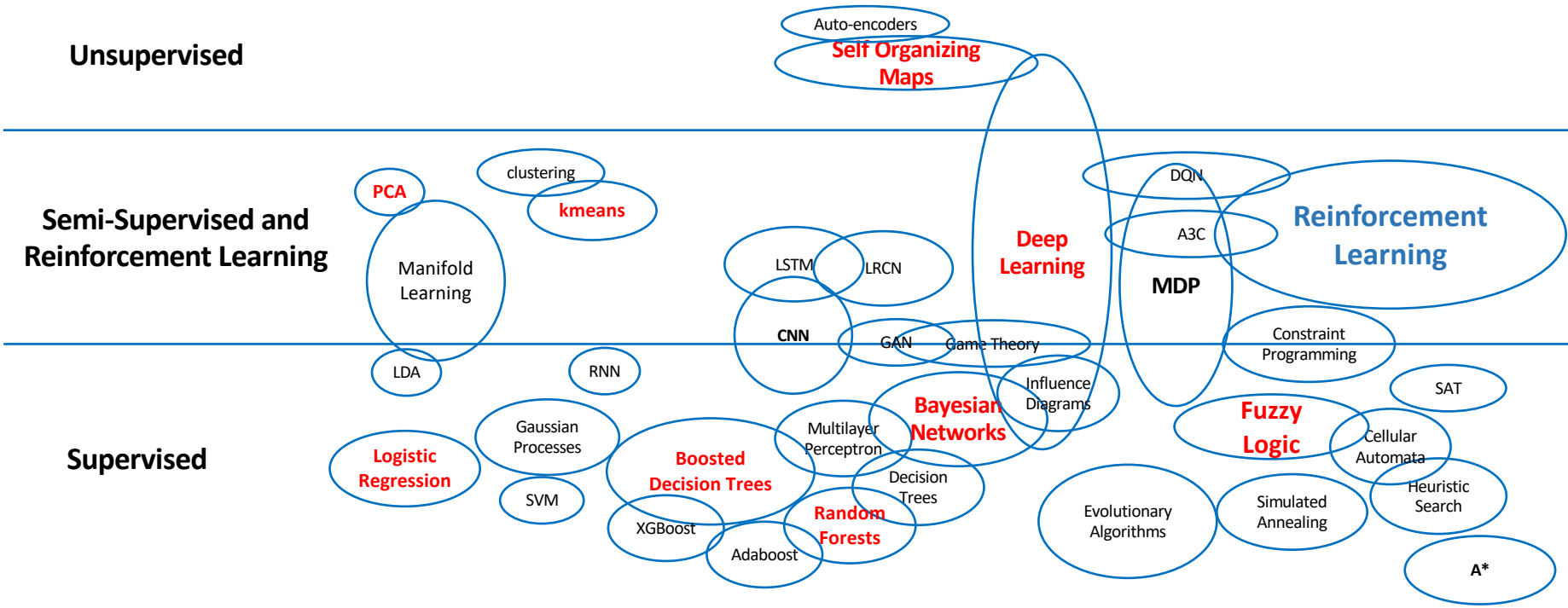
Buzzwords are my jam….

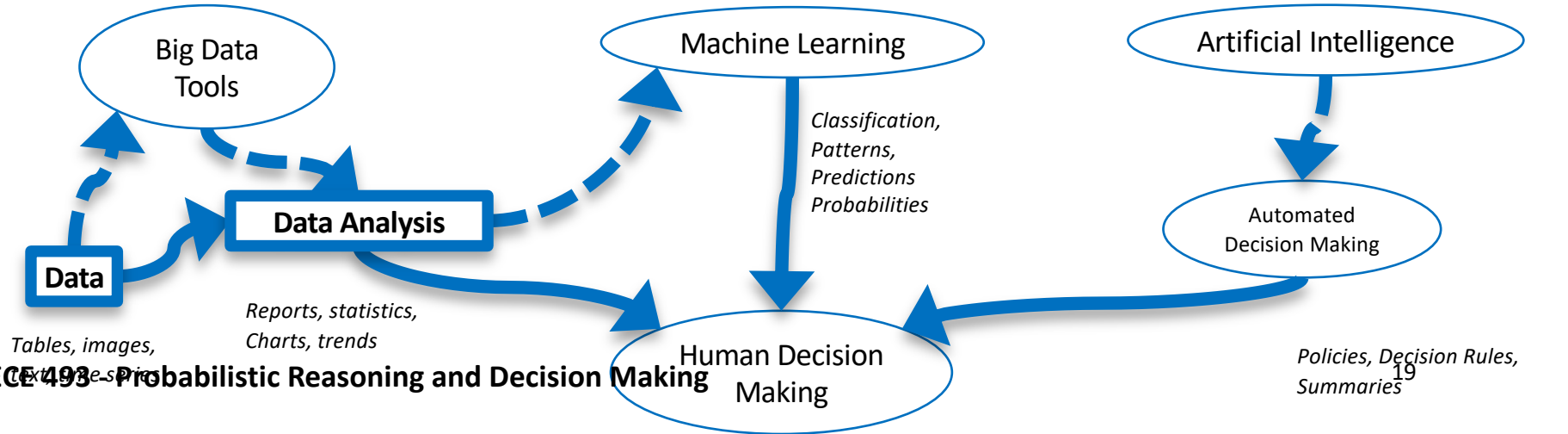
Or "Data, Big Data, Machine Learning, AI, and all that…"



X	f_1	f_2	…	f_k
x_1	.5	104.2		China
x_2	.34	92.0		USA
…				
x_n	.2	85.2		Canada

Or “Data, Big Data, Machine Learning, AI, and all that…”





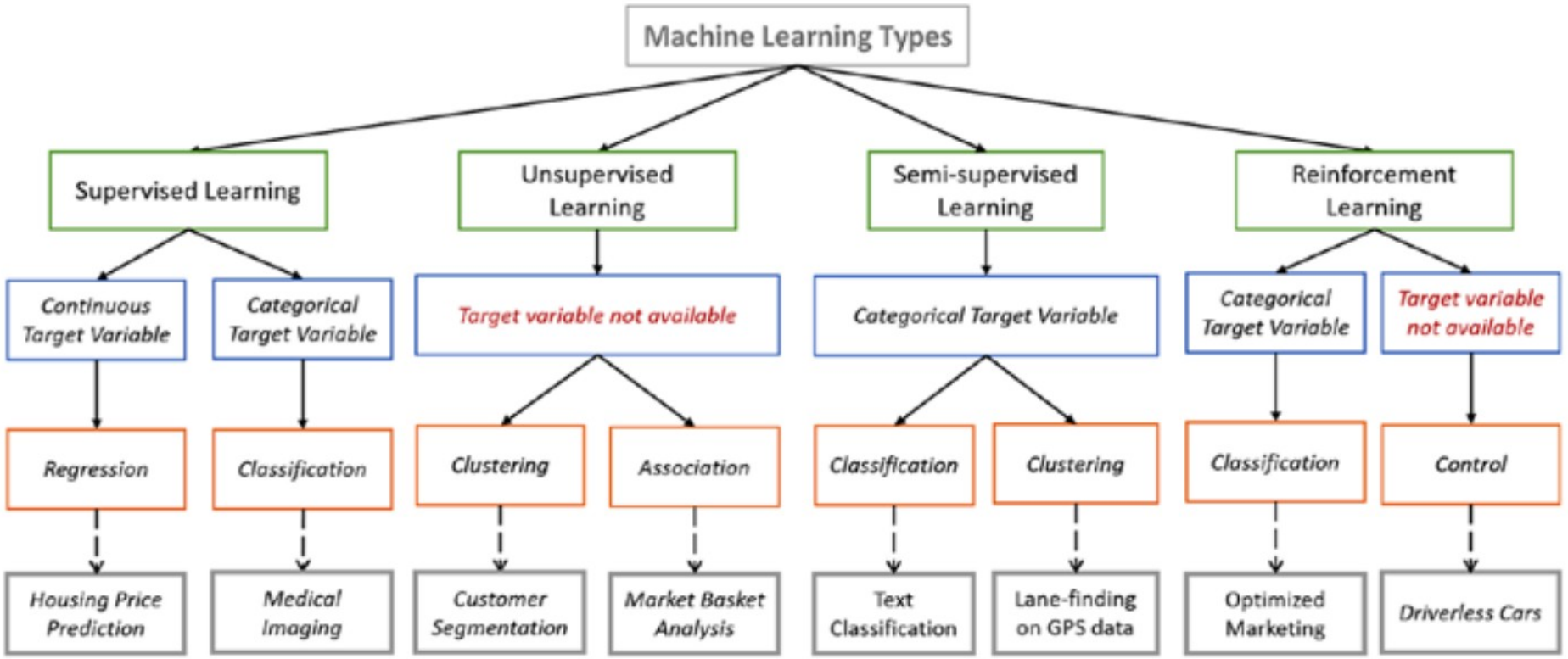
What is Machine Learning?

"Detect patterns in data, use the uncovered patterns to predict future data or other outcomes of interest"

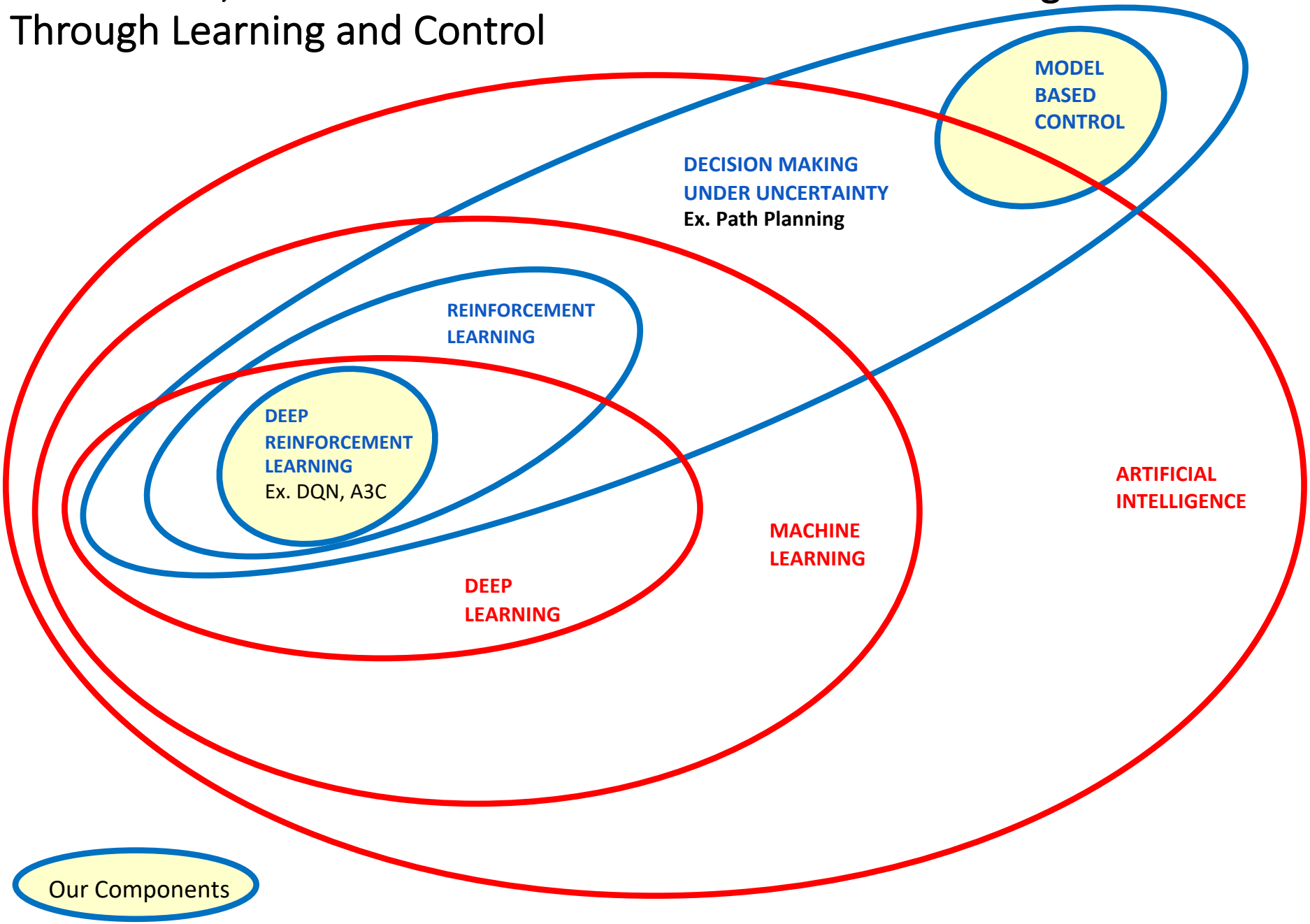
– Kevin Murphy, "Machine Learning: A Probabilistic Perspective", 2012

- Finding patterns in data to allow us to:
  - Group people/events/objects into clusters
  - Classify them according to known meaningful labels
  - Predict their behavior in the future
  - Detect anomalous data points
  - Allow people to make better decisions

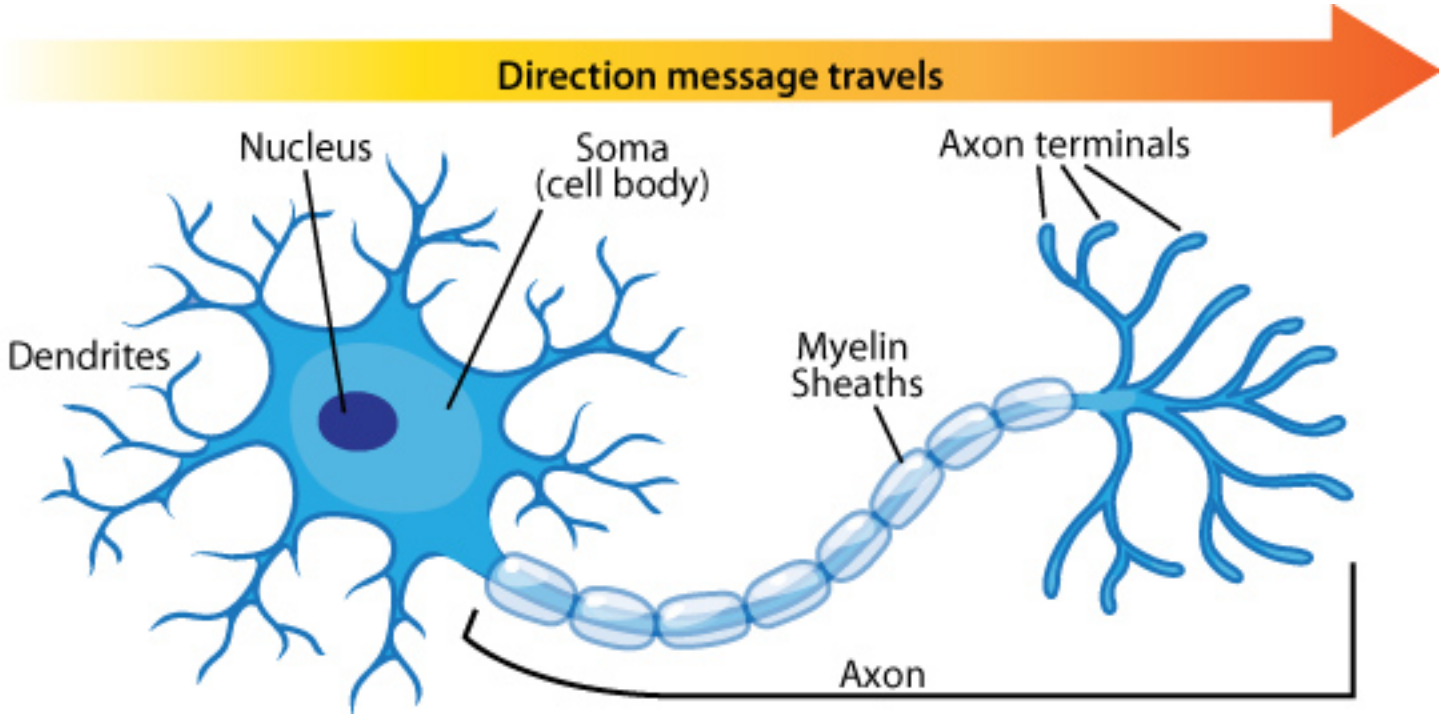
Major Types of Machine Learning



Hierarchical, Collaborative Policies for Autonomous Driving Through Learning and Control



What is a Neural Network?

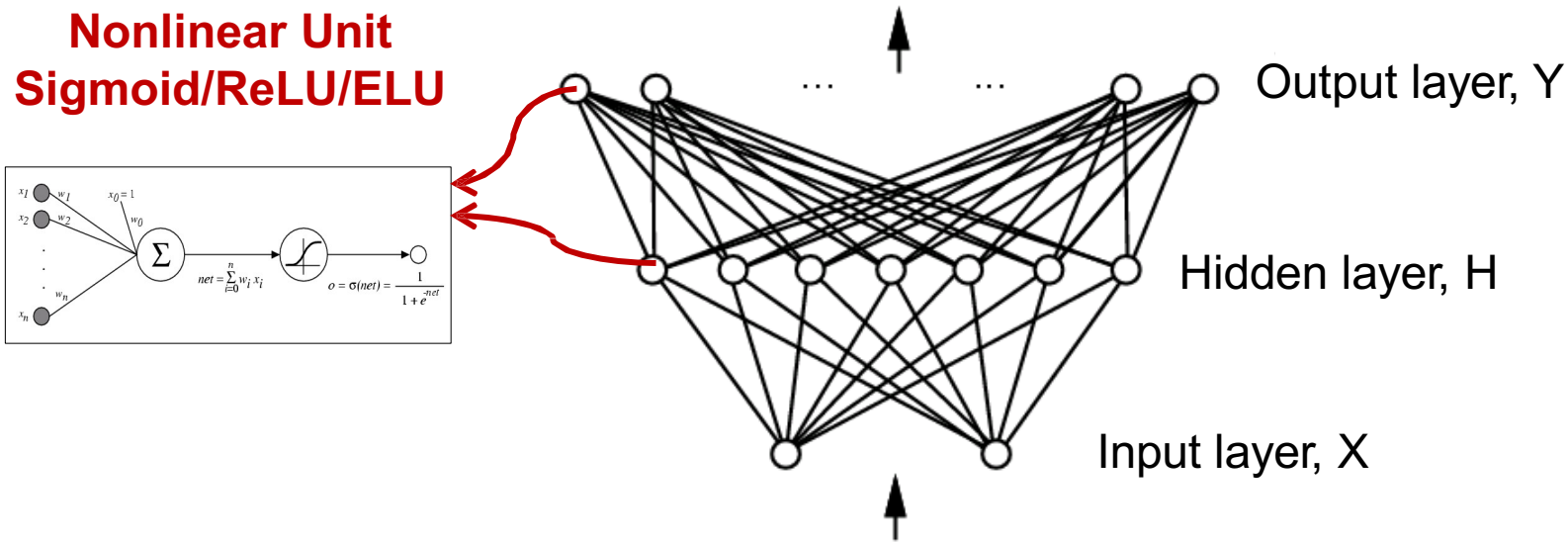


Neural Network and Deep Learning Basics

Neural Networks to learn f : X → Y

- f can be a non-linear function
- **X** (vector of) continuous and/or discrete variables
- **Y** (vector of) continuous and/or discrete variables

Feedforward Neural networks - Represent f by network of non-linear units:



Basic Three Layer Neural Network

Input Layer

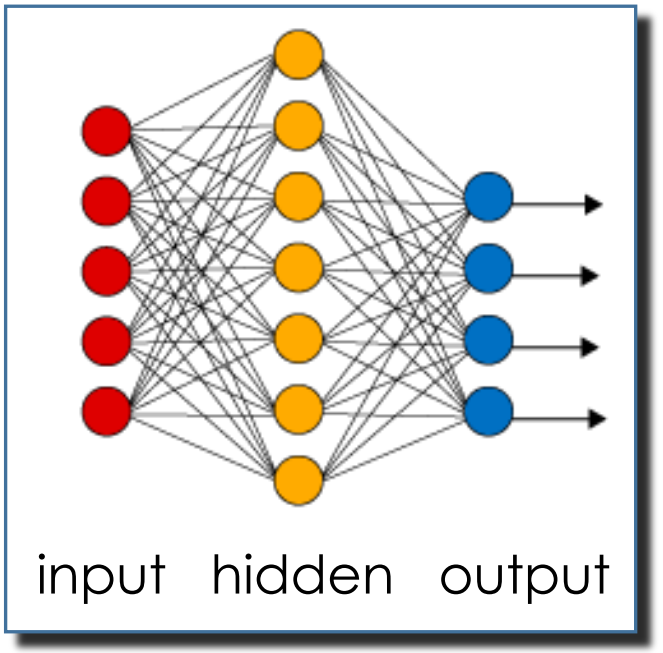
- Takes an array (list) of numbers
- Each input collects one feature/dimension of the data and passes it on to the (first) hidden layer.
- Measurements, aggregate stats, computed indices, pixels

Hidden Layer

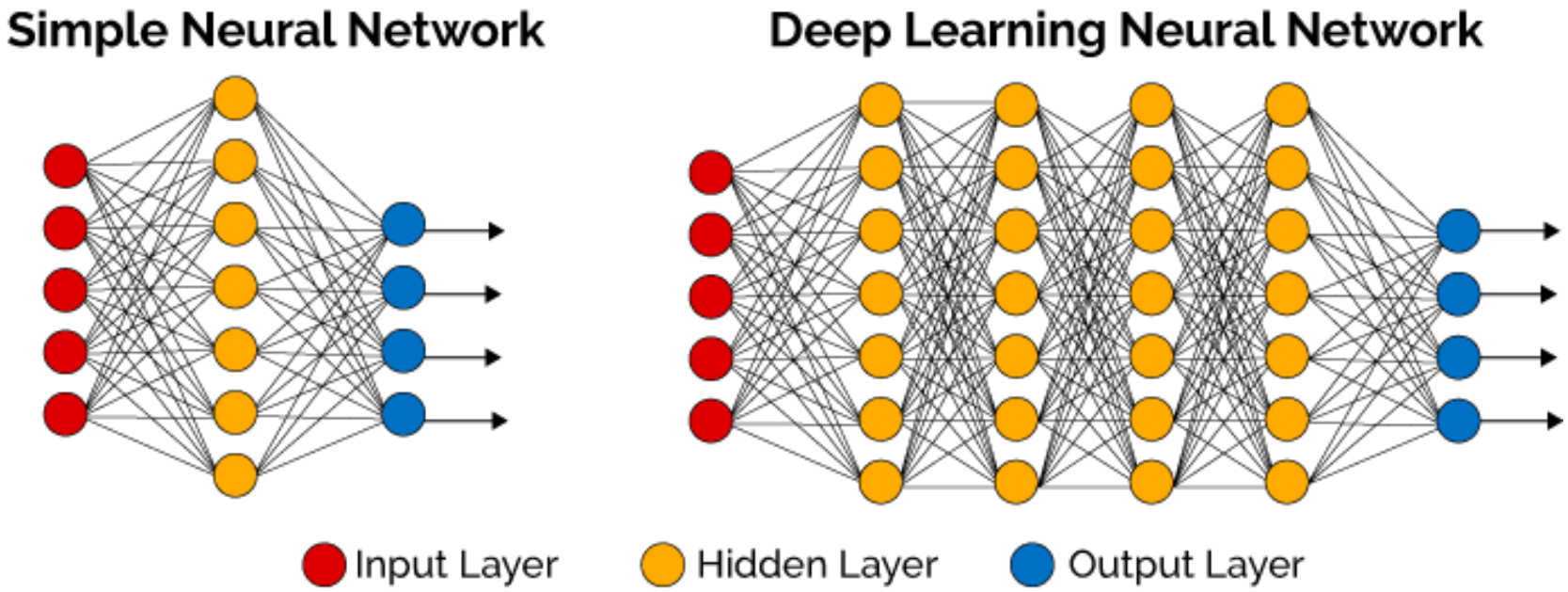
- Each hidden unit computes a weighted sum of all the units from the input layer

Output Layer

- Each output unit computes a weighted sum of all the hidden units and passes it through a threshold function.
- Target variables, predictions, class labels, images

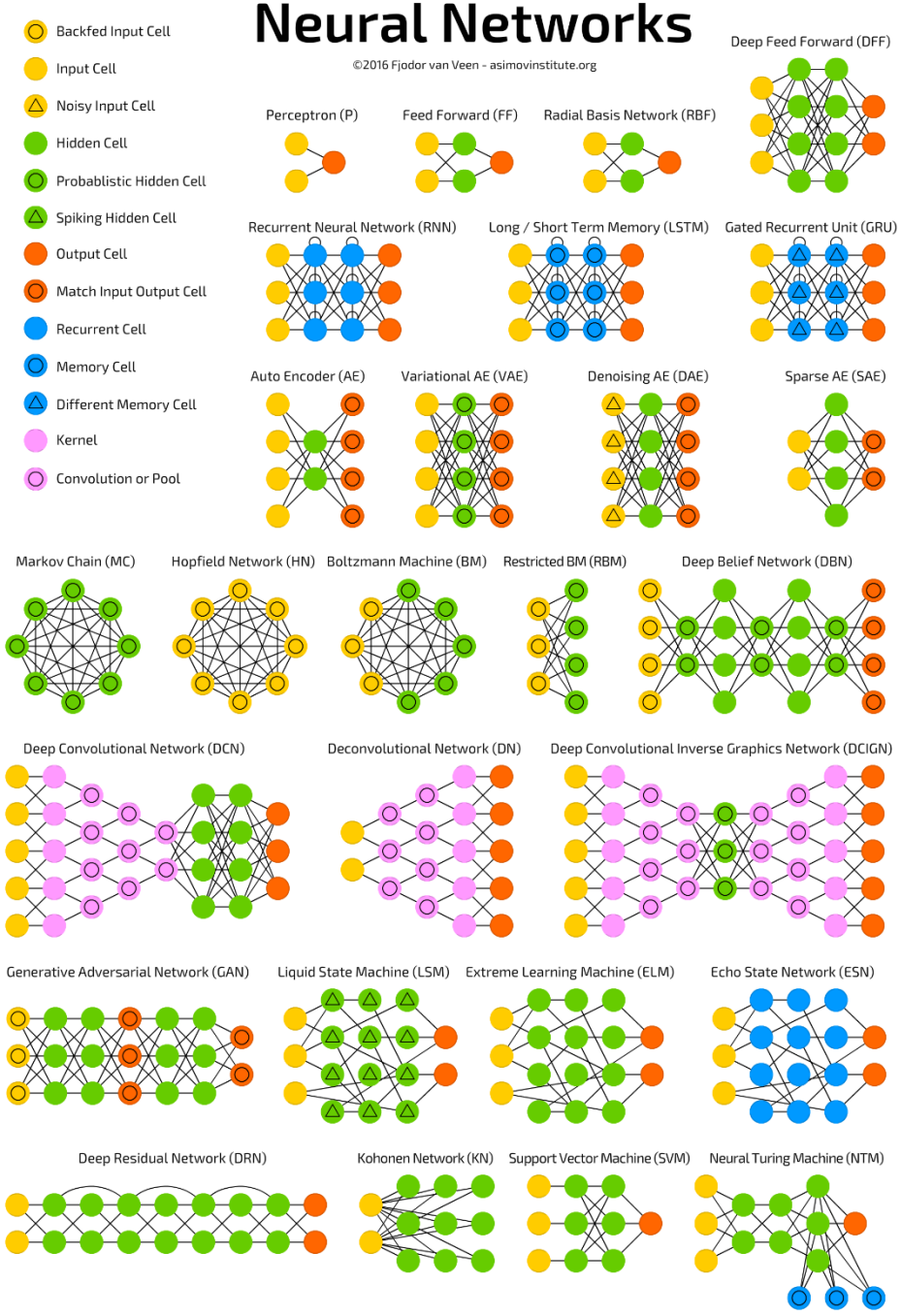


So What is Deep Learning?

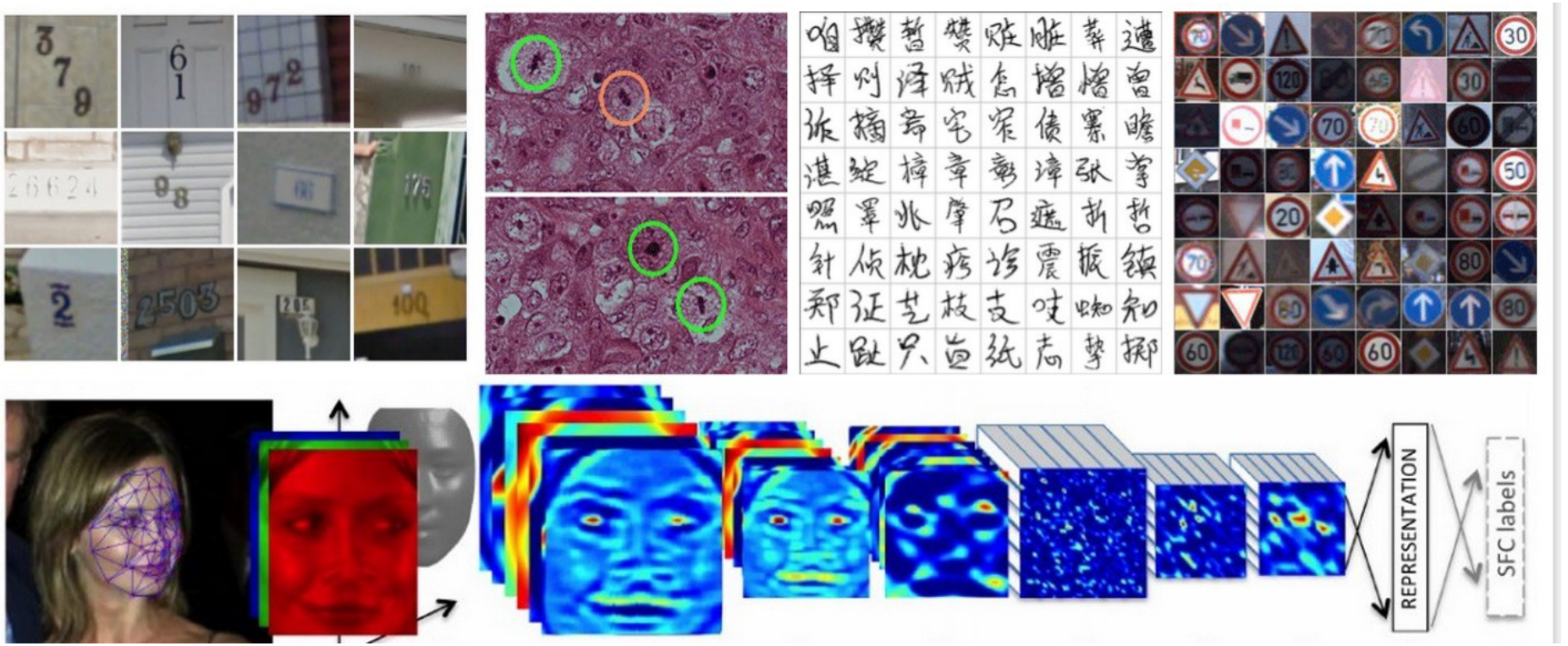


A mostly complete chart of

Neural Networks

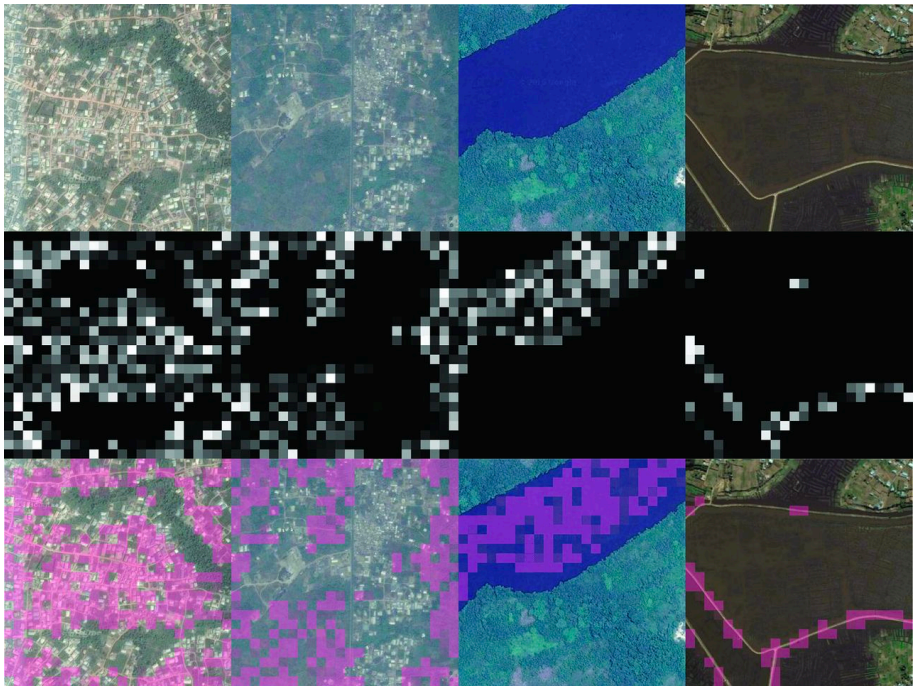


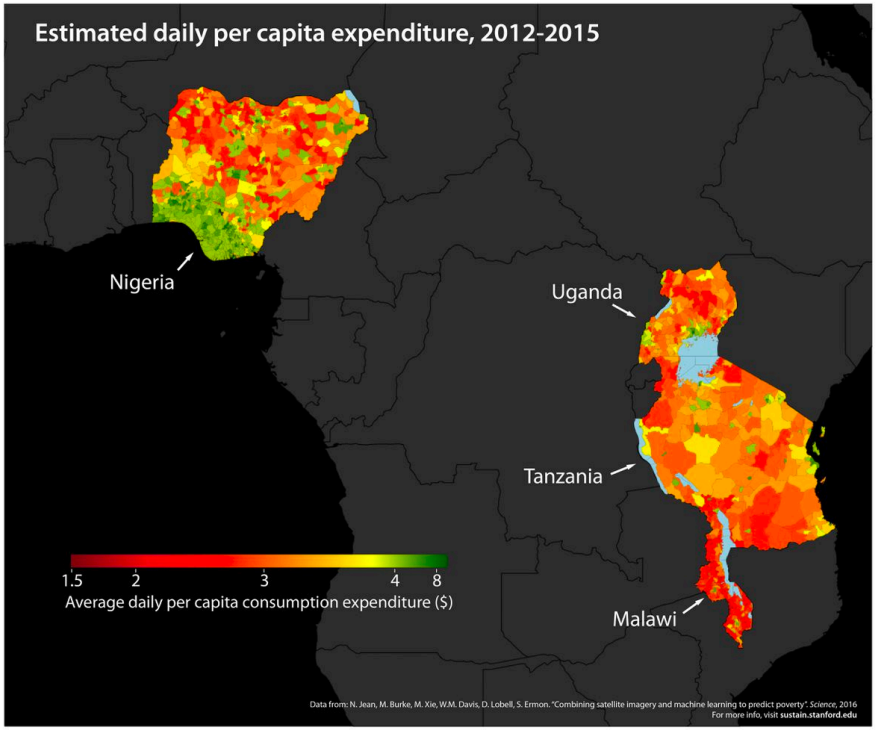
Examples of Convolutional Neural Networks



Using Satellite Images to Predict Poverty Levels

- Stefano Ermon at Stanford uses daytime and nighttime satellite images over Africa to predict poverty levels where data is not available. http://science.sciencemag.org/content/353/6301/790





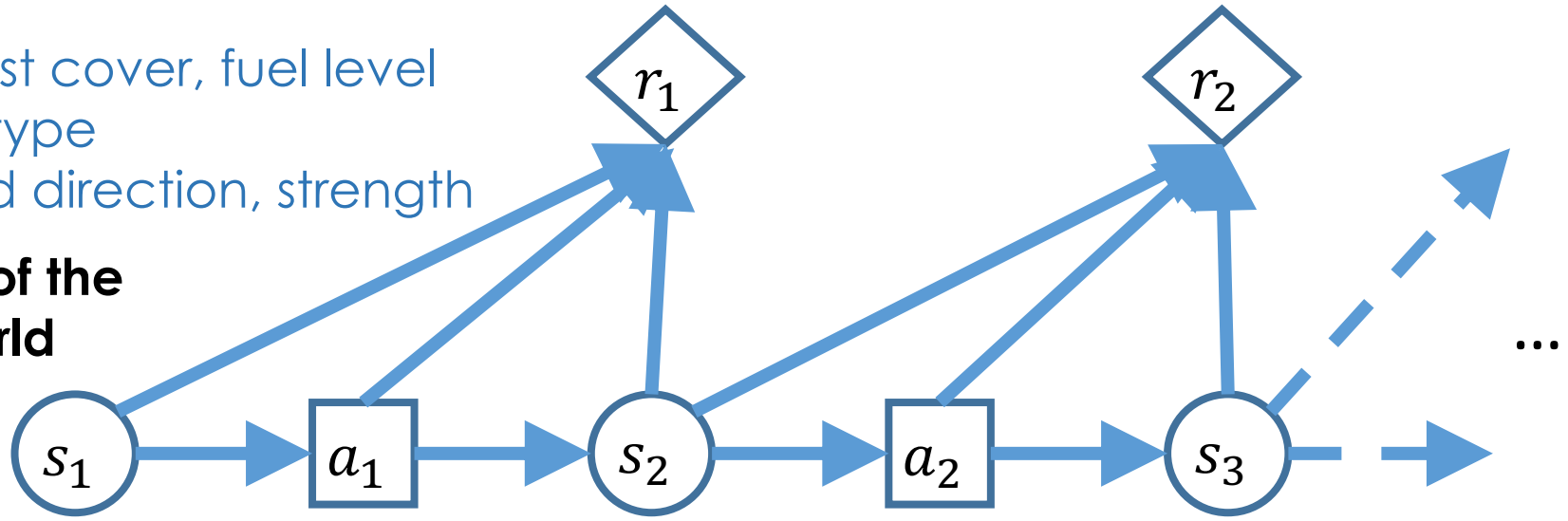
Outline

- A bit about me and my research
- A brief history of AI as I see it
  - Context for the recent Turing Award
  - What does AI mean now?
  - What is Deep Learning?
  - What is Reinforcement Learning?
- Course Outline
- Questions to you about how this course should proceed
  - Lectures vs discussions
  - Reading
  - Collaborative Notes

Markov Decision Process (MDP)

- Fire suppression cost
- Damage, area burned

Rewards



- Forest cover, fuel level
- Soil type
- Wind direction, strength

State of the World

Actions

- Let Burn/Suppress
- Fireline (x,y,shape,…)

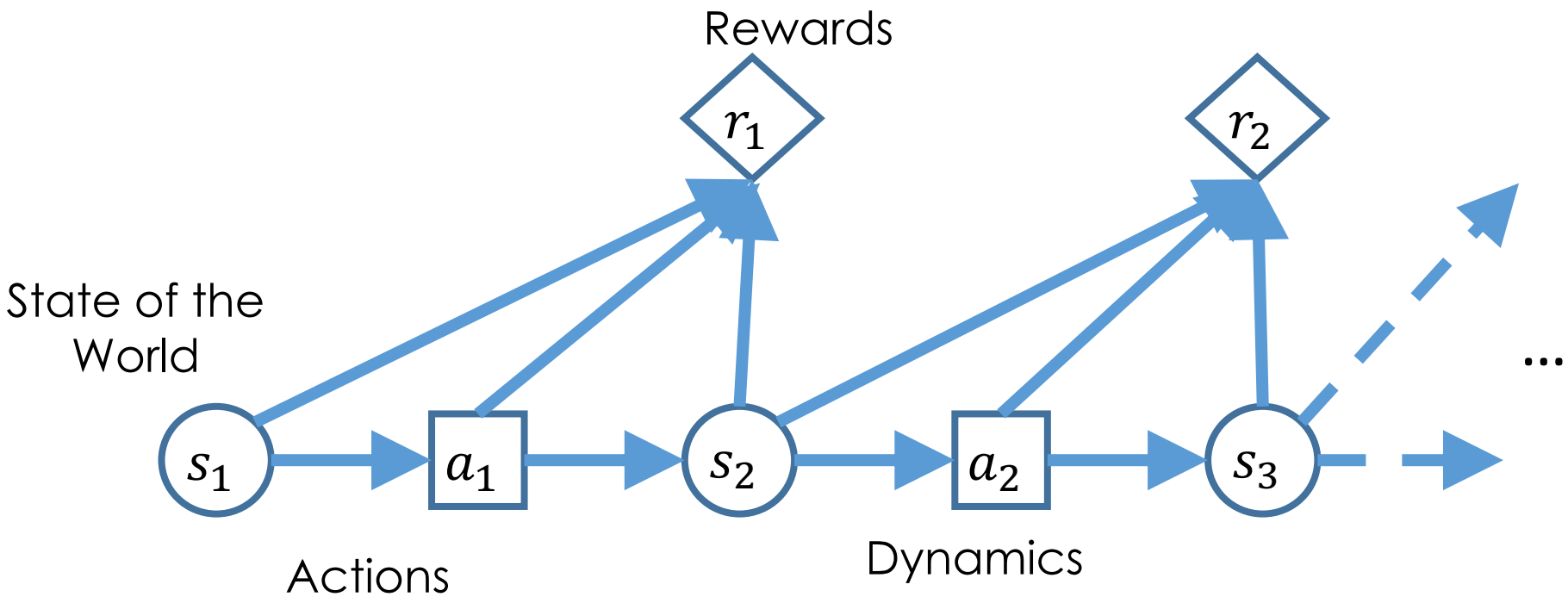
Dynamics

- Can be deterministic or random
- Fire spread
- Upcoming Weather
- Fuel/soil cycles
- Human behaviour

Markov Decision Process (MDP)

Different Fields of AI Come Down to

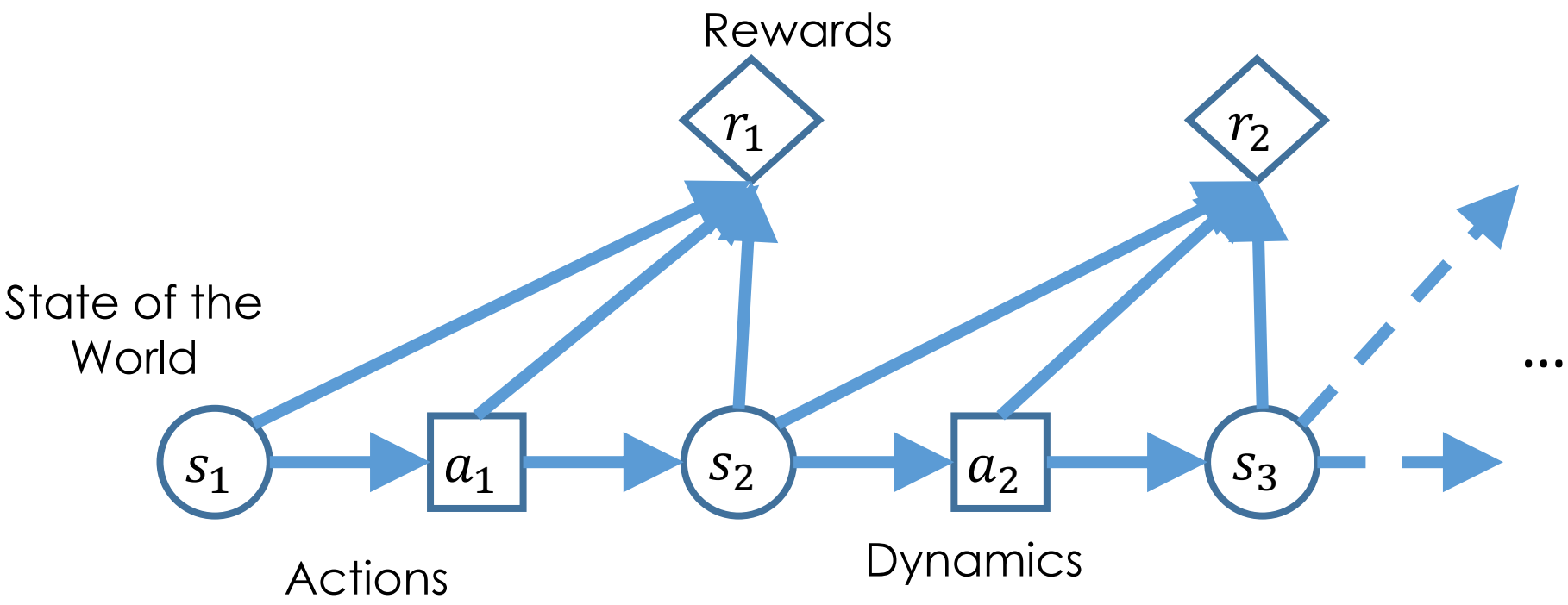
- Which part of this picture do you know?
- Which can you estimate?
- Which do you *need* to know?



Reinforcement Learning (RL)

In this field we do not know ahead of time

- The rewards (the relative value of outcomes)
- The dynamics (how the world works)
- But…we can *ask* for them by acting



Reinforcement Learning as an MDP

Reinforcement Learning is learning the policy for taking actions for an MDP when you **do not have access** to the full definition of:

- the rewards
- AND/OR the dynamics

Training must be carried out interactively :

1. Commit to action using latest (or some) policy
2. Find out the next state and reward from the world/simulator/environment
3. Improve your policy
4. Repeat until the policy is "good enough" or it stops changing

The "Physics" of Reinforcement Learning

- RL always comes down to solving a recursive Bellman Equation that relates the values of states and actions in those states.
- Comes in many forms, MDP can be solved exactly, but only in small cases with complete knowledge.
- RL algorithms seek seek to iteratively updating value function, or the policy directly, through experience to make improved decision decisions

Value Iteration

$$V^*(s) = R(s) + max_a \gamma \sum_{s'} P(s'|s,a) V^* s'$$

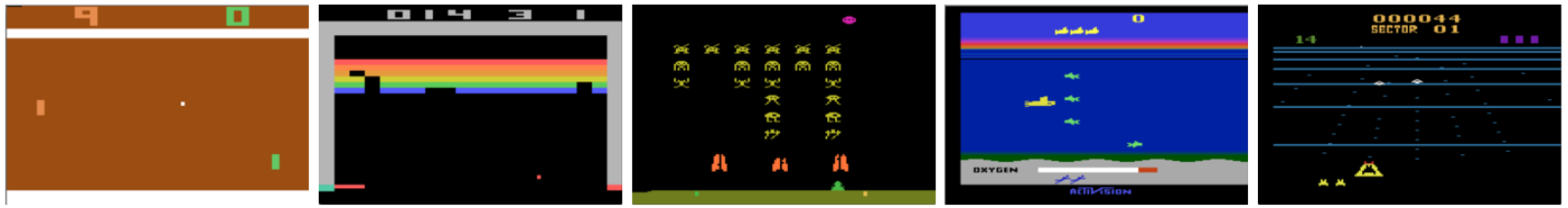
Policy Gradient

$$\nabla_\theta V^\pi(\mathbf{s}_0) \approx \frac{1}{|K|} \sum_{\mathbf{k} \in K} R(\mathbf{k}) \sum_t \nabla_\theta \log \pi(\mathbf{a}^{\mathbf{k},t} | \mathbf{s}^{\mathbf{k},t}, \theta)$$

Q-learning

$$Q'(s_t, a_t) = (1 - \alpha) Q(s_t, a_t) + \alpha (r_t + \gamma \max_a Q(s_{t+1}, a))$$

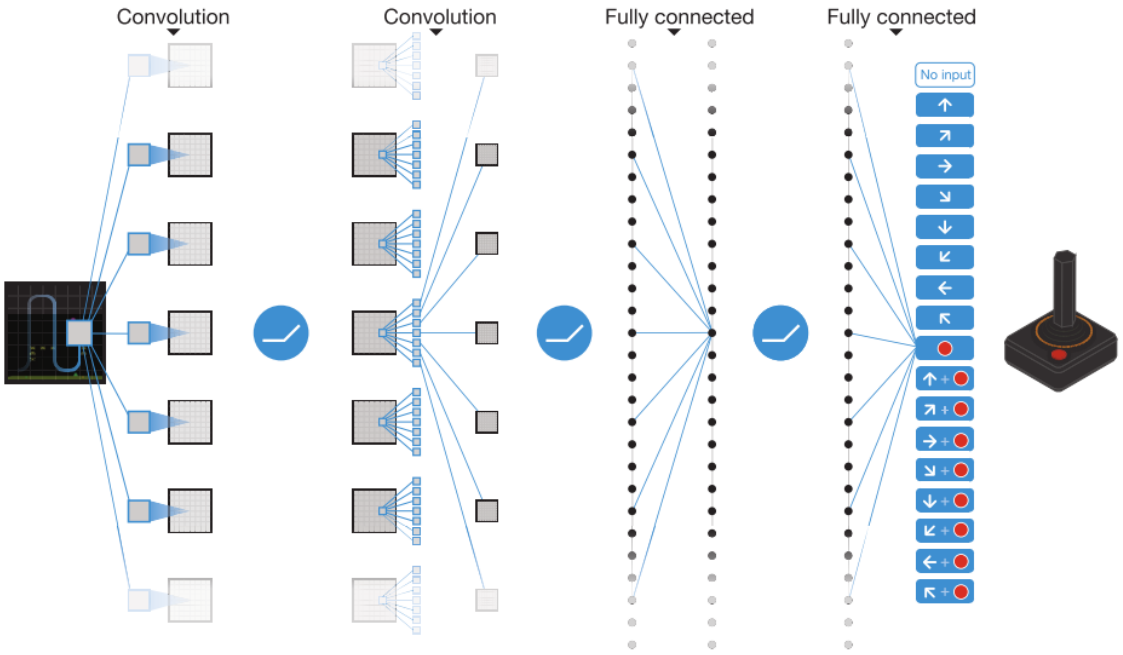
Deep Reinforcement Learning on Atari Games



Flurry of advances since 2014 by Google DeepMind and others applying Deep Learning to RL algorithms.

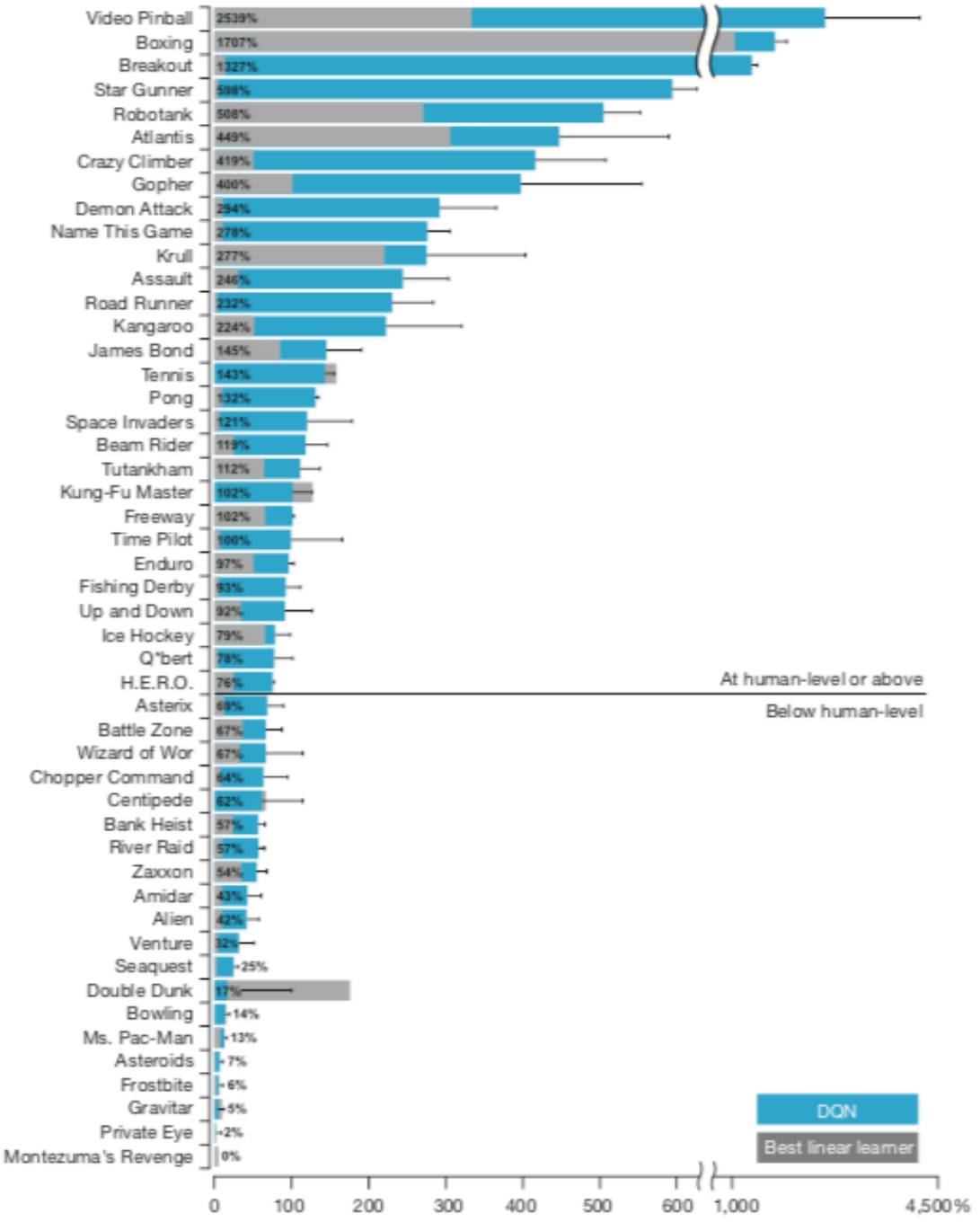
Many algorithms since then trying to provide a better way to learn the value function with DNNs

- Alpha Go – RL + human training
- Alpha Zero – RL + MCTS search + playing itself (Go, Chess)

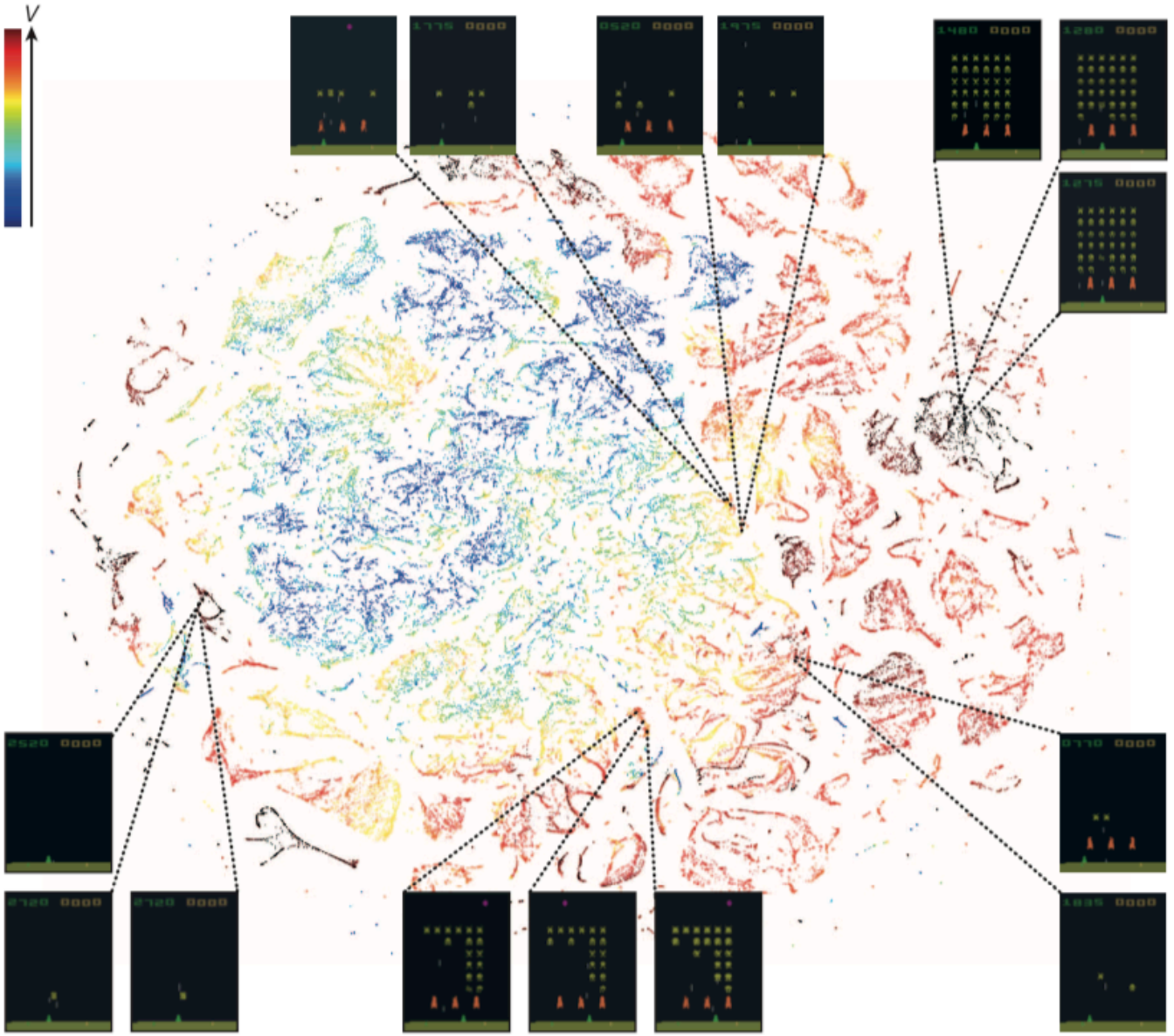


Initial performance of DQN against Atari

Note the line of human-level performance

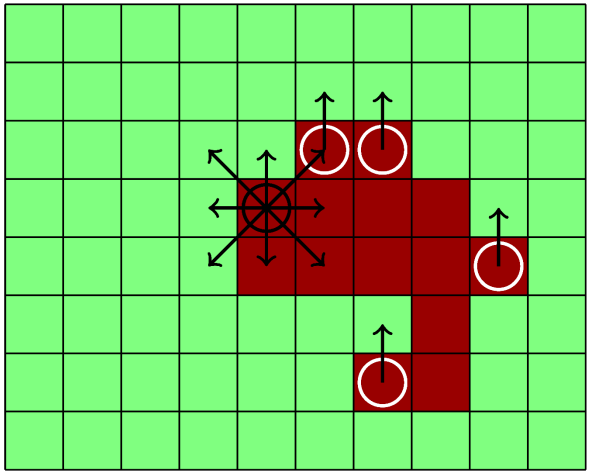


But what is it “thinking” about?

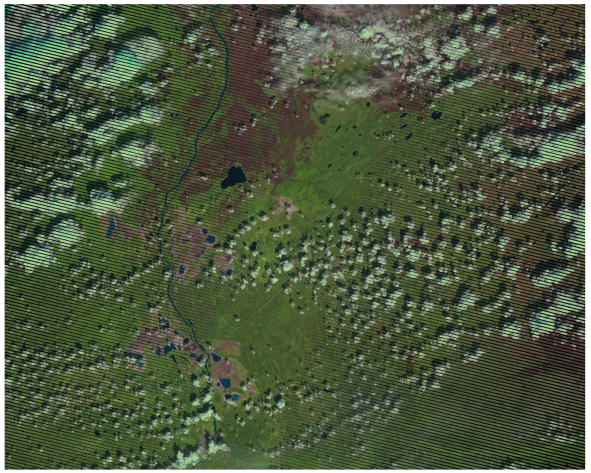


Problem setup

- Using satellite images from two large forest wildfires in Northern Alberta : Richardson 2011, For McMurray 2016.
- Used just publically accessible low rez images from USGS, (Landsat)



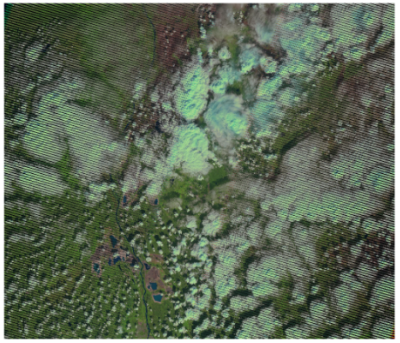
(a) Schematic of the state and actions



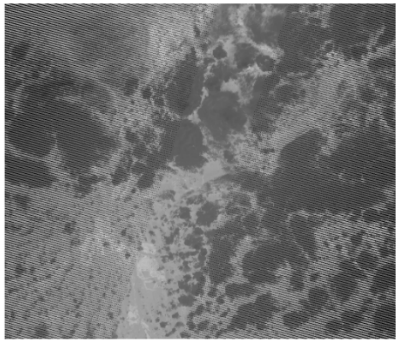
(b) Raw Color Image

(c) Thermal Image

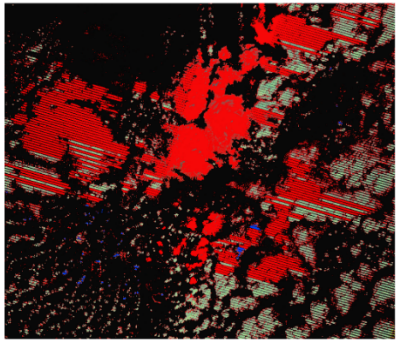
Forest Fire Prediction – Results



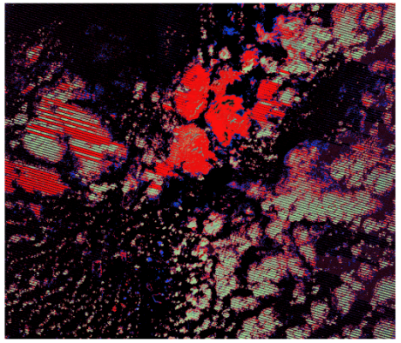
(a) Satellite Image of August 11



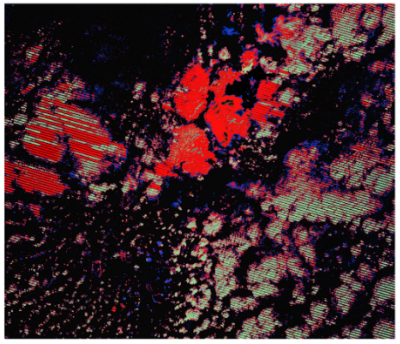
(b) Thermal Image of August 11



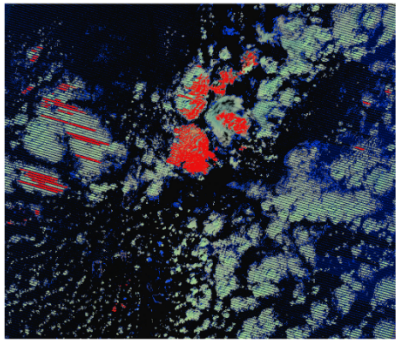
(c) Gaussian Process



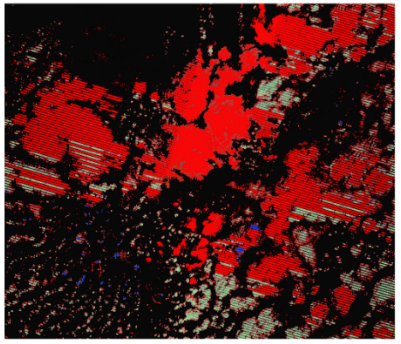
(d) Value Iteration



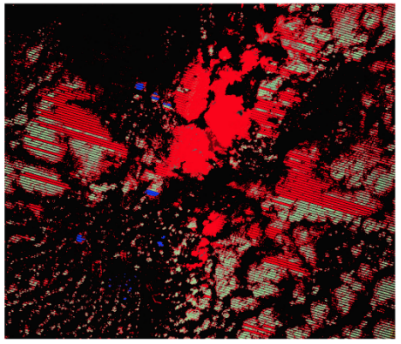
(e) Policy Iteration



(f) Q Learning



(g) MCTS



(h) A3C

Prediction of Spatially Spreading Processes

Using Forest Wildfires as a demonstration domain:

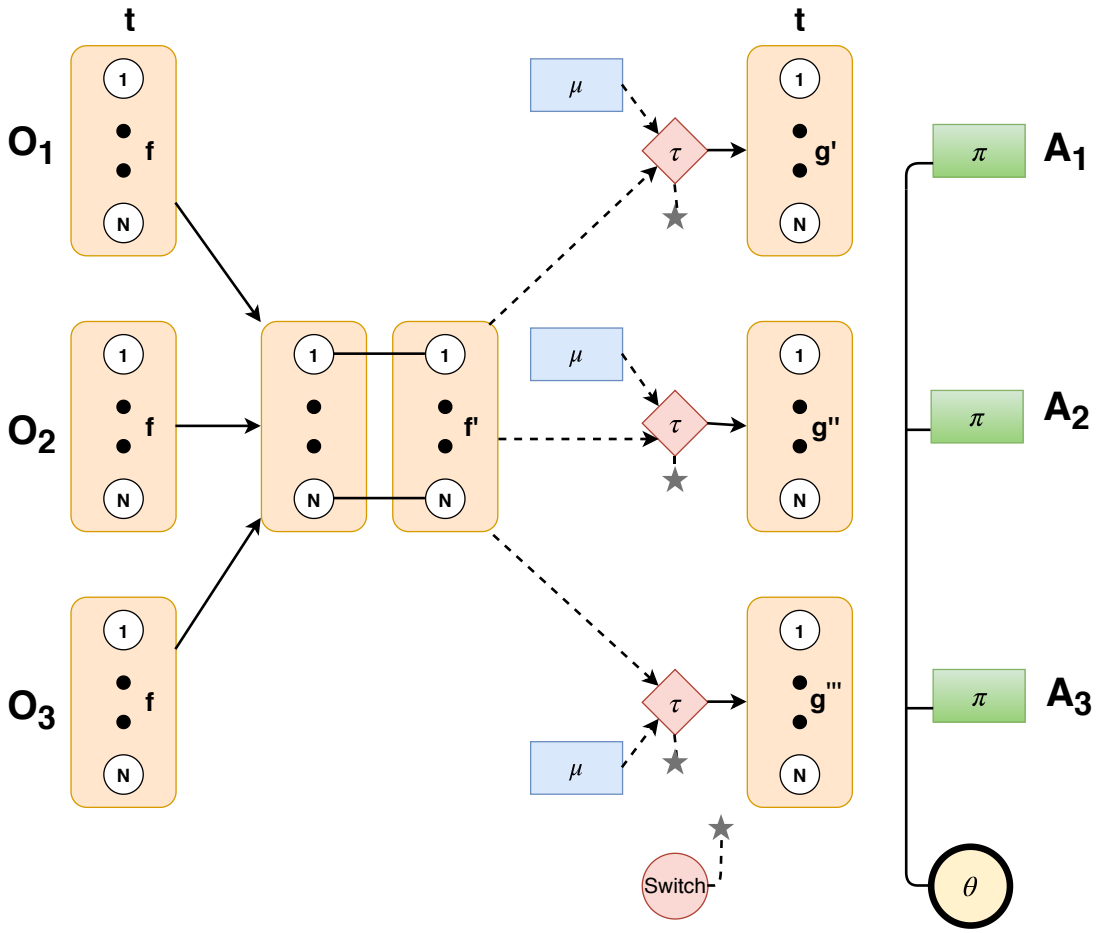
- **RLDM 2017** - The idea **of using RL for learning dynamics from image data**, comparison of classical RL algorithms with DQN [Subramanian and Crowley (2017)]
- **Frontiers in ICT Journal 2018** - Compared to broader range of RL methods Also looked at Gaussian Processes for a fully supervised comparison Noticed a **tradeoff** between MCTS and A3C [Ganapathi Subramanian and Crowley (2018)]
- **CAI Paper 2018** - proposed a new algorithm MCTS-A3C to **take advantage of strengths of both**.

Goal: Demonstrate that learning an agent-based style model be done from raw data via RL, and can provide comparable results to other methods.

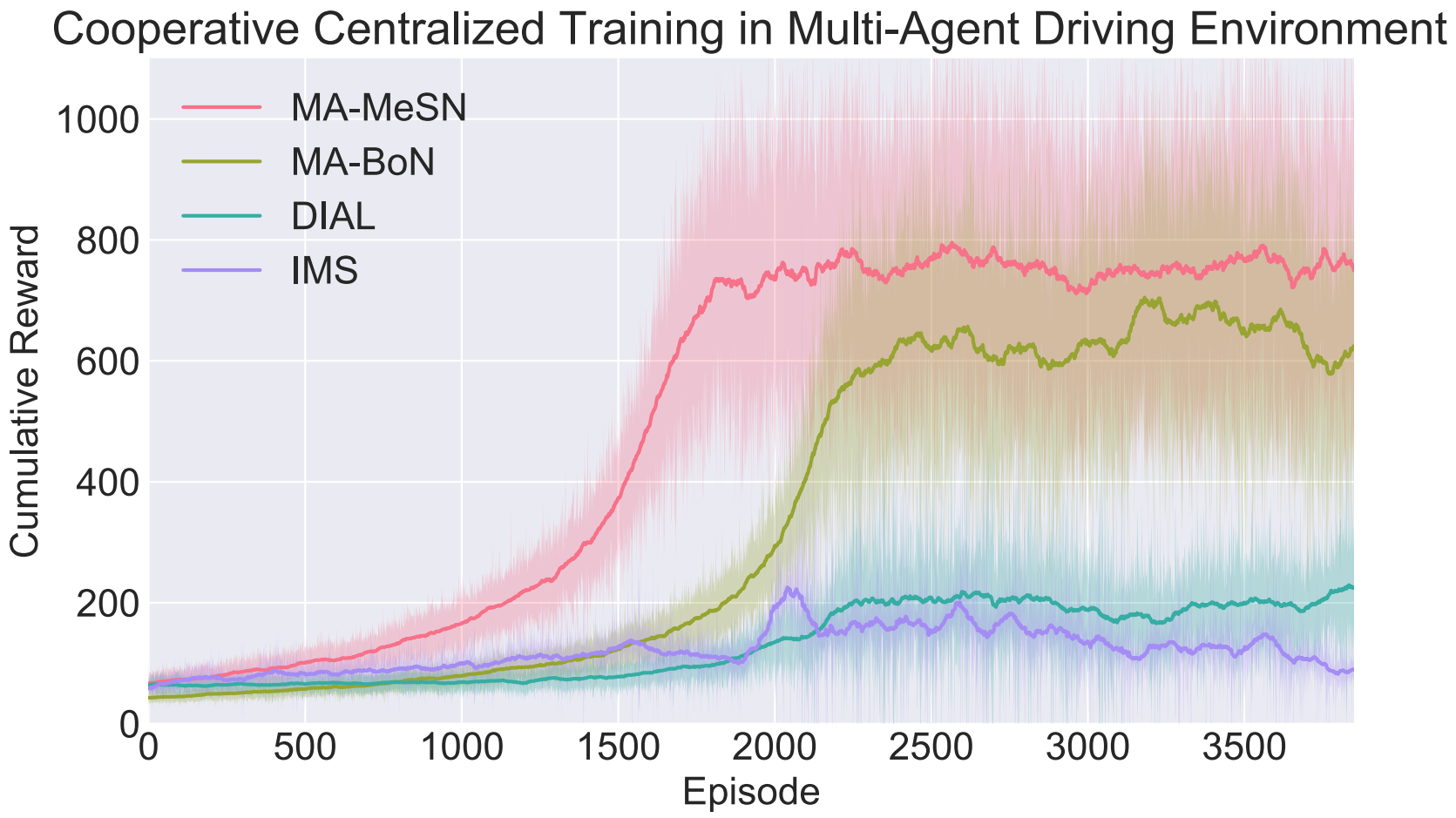
Multi-Agent Broadcast Network (MA-BoN)

An extension to IMS to achieve a cooperative MARL policy for homogeneous agents with a reduced communication channel throughput

We replace the averaging module with a network trained to generate a **single broadcast message** (vs multiple rounds of messages in IMS)



Results


Cooperative Centralized Training in Multi-Agent Driving Environment
MA-MeSN
MA-BoN
DIAL
IMS
Cumulative Reward
1000
800
600
400
200
0
0
500
1000
1500
2000
2500
3000
3500
Episode

Or “Data, Big Data, Machine Learning, AI, and all that…”