Automated Material Synthesis using Deep Reinforcement Learning

Mark Crowley

Assistant Professor

Electrical and Computer Engineering

University of Waterloo

Waterloo, Ontario, Canada

https://uwaterloo.ca/scholar/mcrowley

@compthink

UNIVERSITY OF WATERLOO

WATERLOO.AI

WATERLOO ARTIFICIAL INTELLIGENCE INSTITUTE

ELEMENT AI

About Me

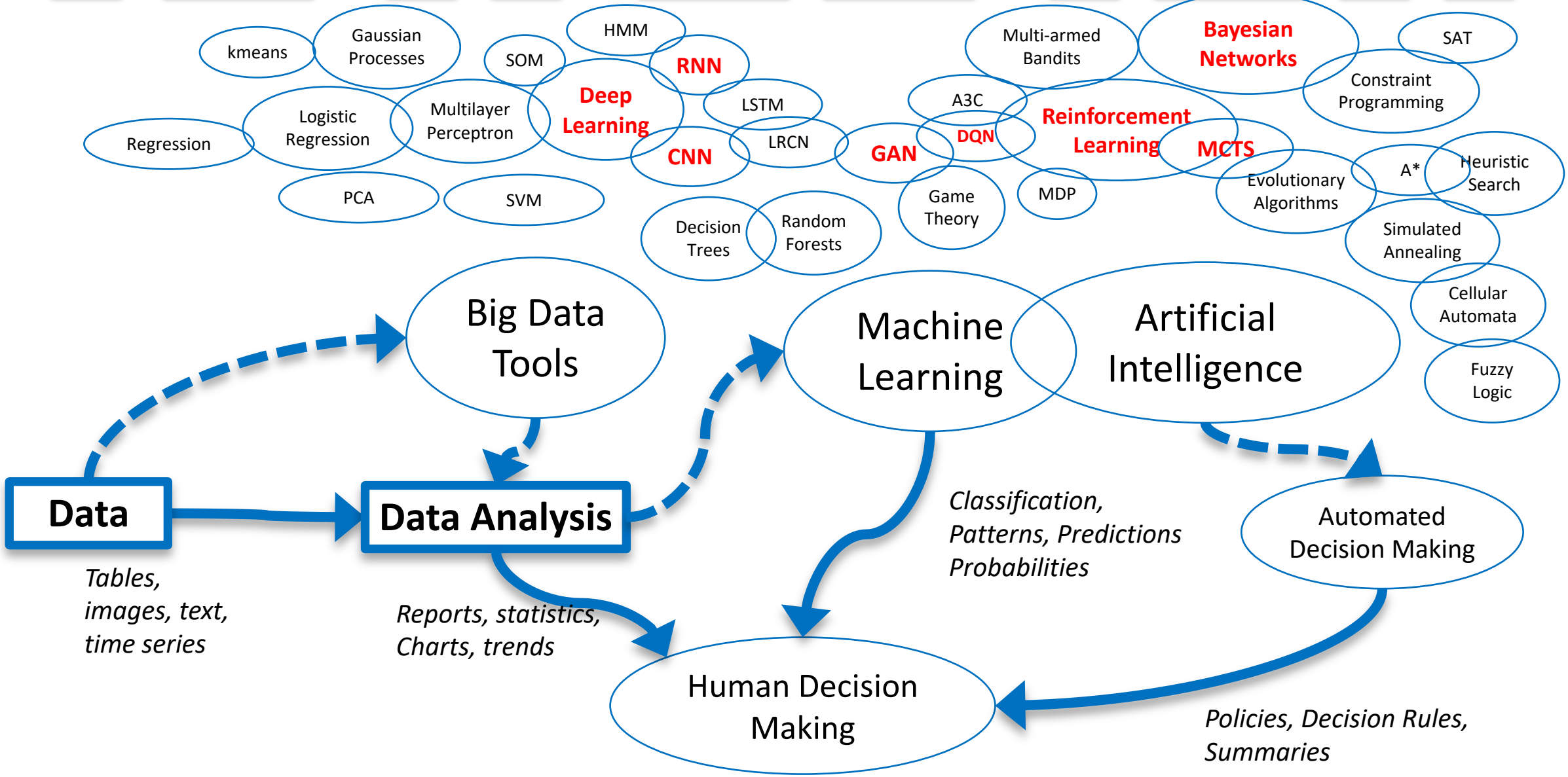
- B.A. : York University Computer Science (1999)
- PhD : Computer Science at the University of British Columbia (2011)
- Postdoc : Oregon State University
  - on Computational Sustainability
- Assistant Professor (since 2015)
  - University of Waterloo – ECE Department
  - UW ECE Machine Learning Lab
    - (https://uwaterloo.ca/scholar/mcrowley/lab)
  - Waterloo.ai : Waterloo Artificial Intelligence Institute (http://waterloo.ai)
  - Element AI: Faculty Research Fellow

Multi-Step Chemistry as an AI Problem

i.e. using AI to build an automated chemist advisor

- Can we utilize Artificial Intelligence to automate aspects of physical chemistry to discover new pathways for creating materials with desired properties?
- Artificial Intelligence, Machine Learning and all that…

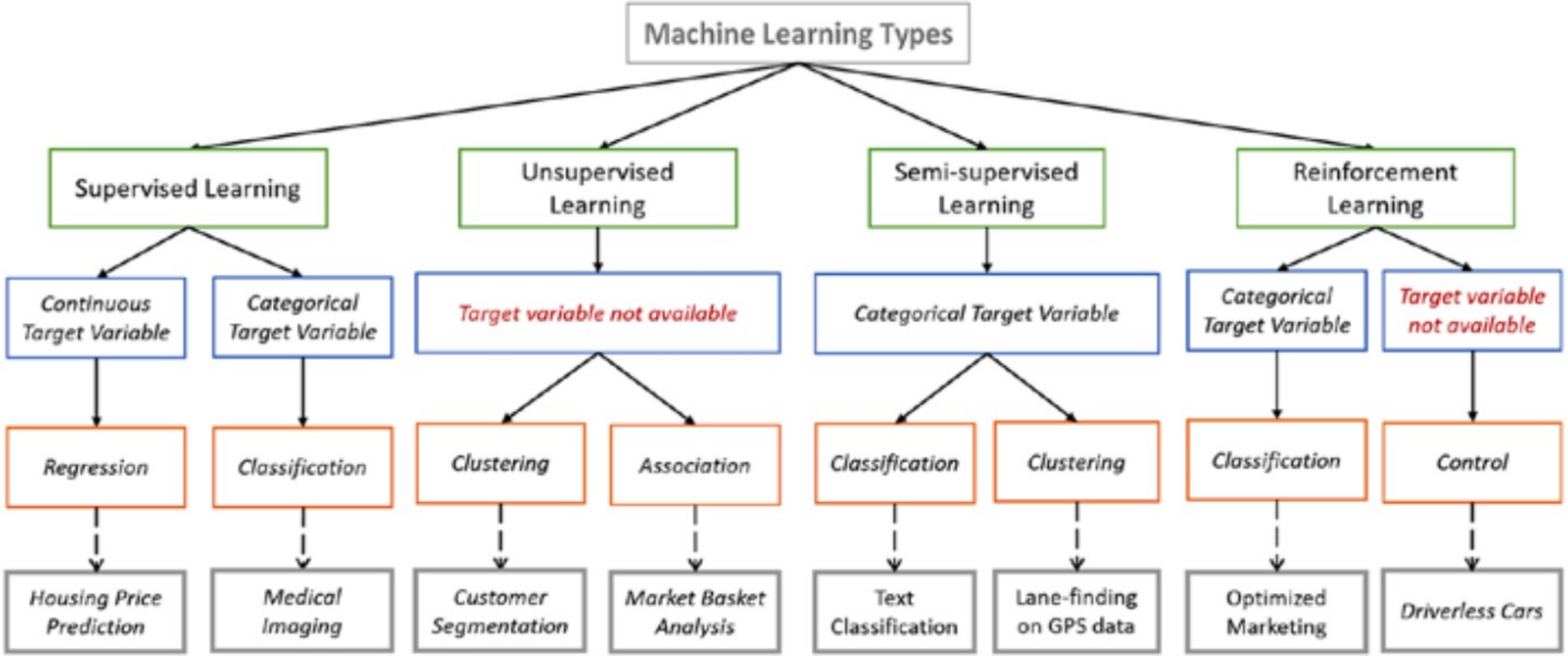
Data, Big Data, Machine Learning, AI, etc, etc,…



Major Types of 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



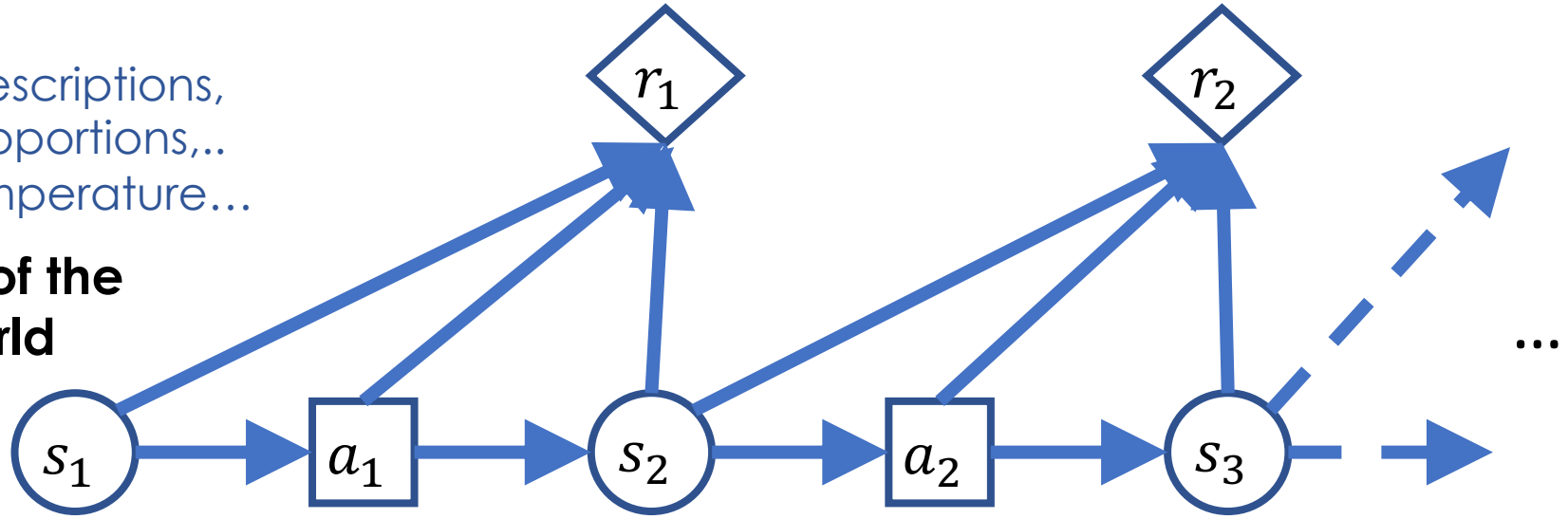
Markov Decision Process (MDP)

- Amount of desired material
- Time spent
- Cost of inputs

Rewards

- Chemical descriptions, amounts, proportions,..
- Pressure, Temperature…

State of the World



Actions

- **Add/remove** reactant
- **Change** temperature/pressure
- **Choose** which "bench" to use
- **Measure** something about your current samples

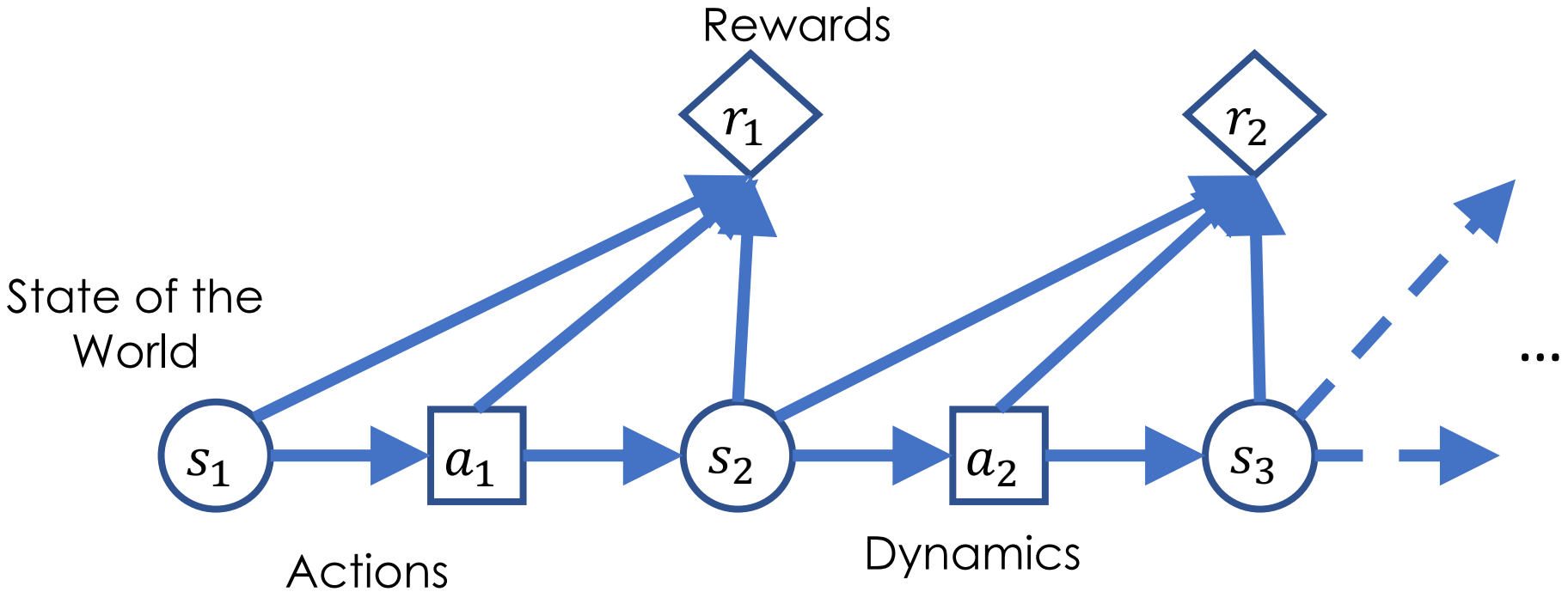
Dynamics

- Well understood but expensive ODE's describing reactions
- Statistical behavior of extraction, outcome chemicals, temperature, pressure

Markov Decision Process (MDP)

Different Fields of AI Come Down to (one way to look at it)

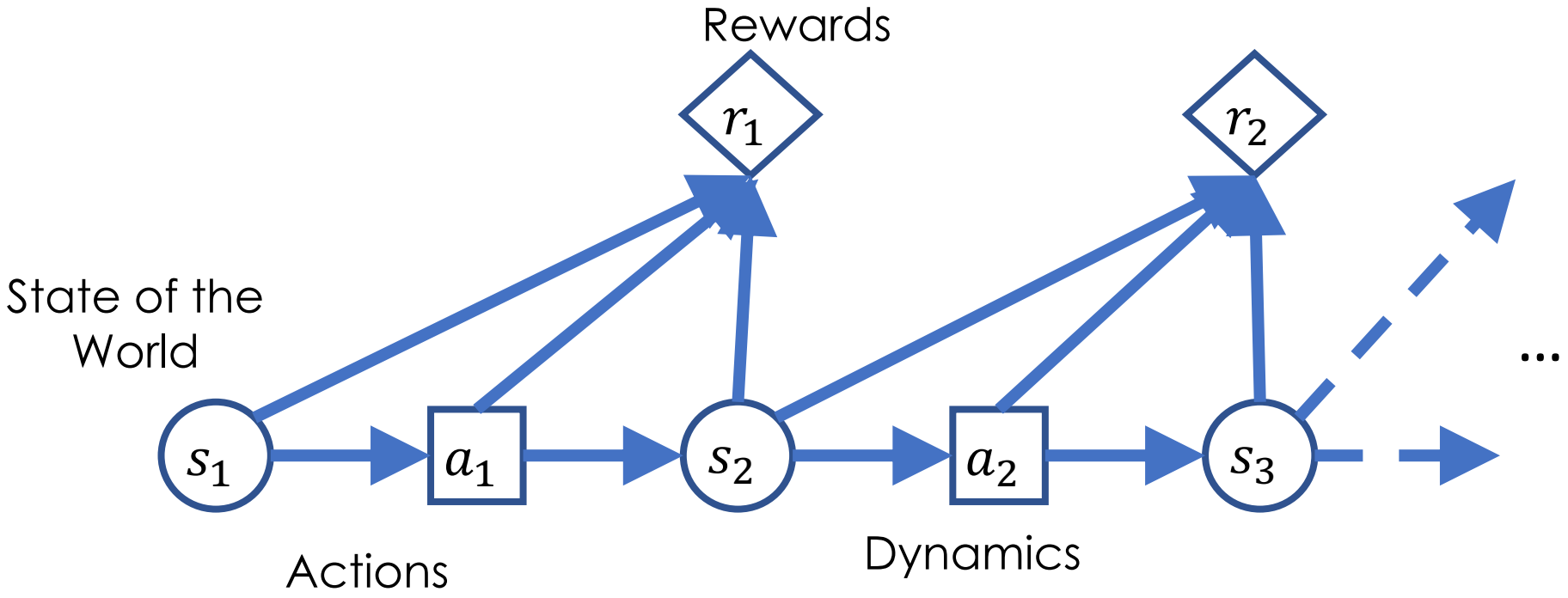
- Which parts of this picture do you know?
- Which can you estimate?
- Which do you *need* to know?



Reinforcement Learning (RL)

In this field what we *do* know ahead of time…

- The (simplified) dynamics for basic bench activities
- The immediate costs of each activity, and our distance to the final goal

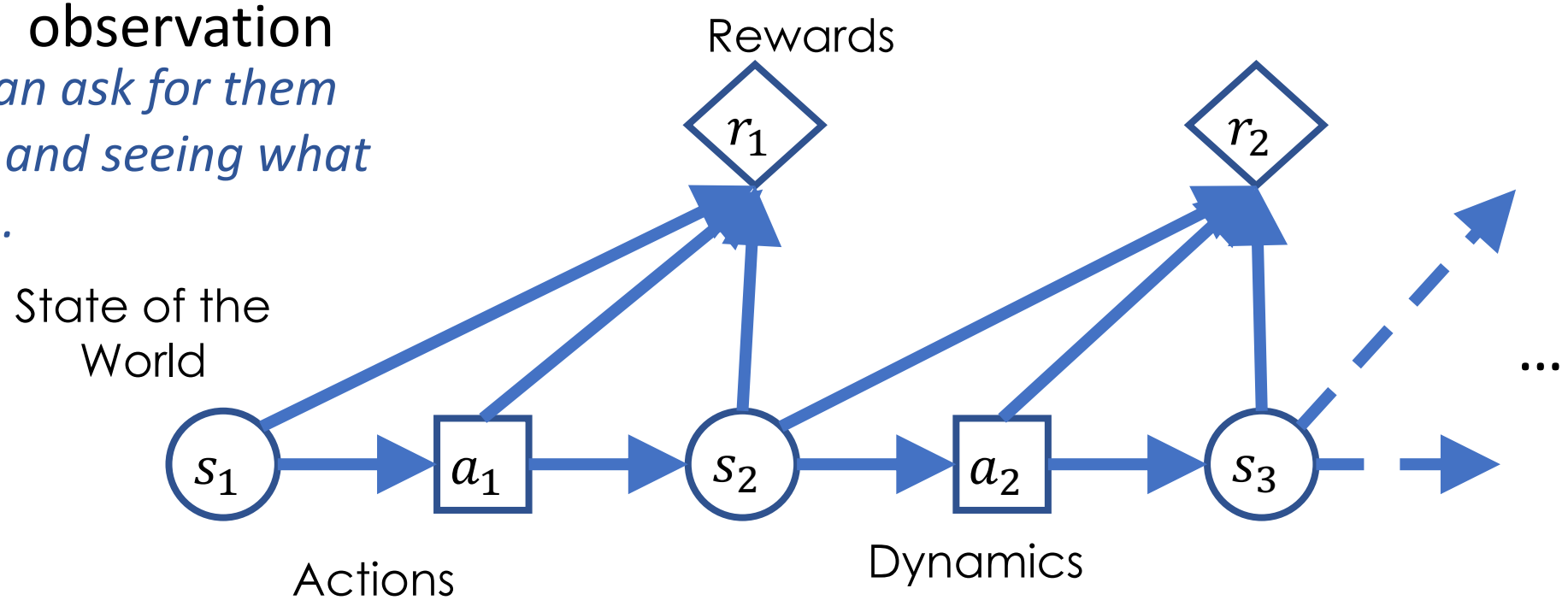


Reinforcement Learning (RL)

But we *do not* know ahead of time…

- The best (or any) way of stringing together a *series* of transformation activities to achieve a desired material
- The full state of the output of an activity without a destructive observation

But, we can ask for them by acting and seeing what happens…



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**
- This in many varieties and it **usually solved approximately**
  - MDPs can be solved exactly, but only in small cases with complete knowledge.
- RL algorithms seek to iteratively update a value function, or the policy directly, through experience to make improved decision decisions

Value Iteration

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

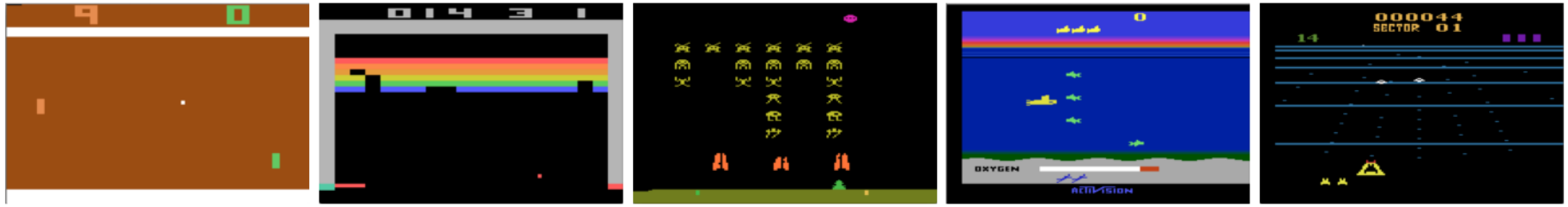
Policy Gradient

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

Q-learning

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

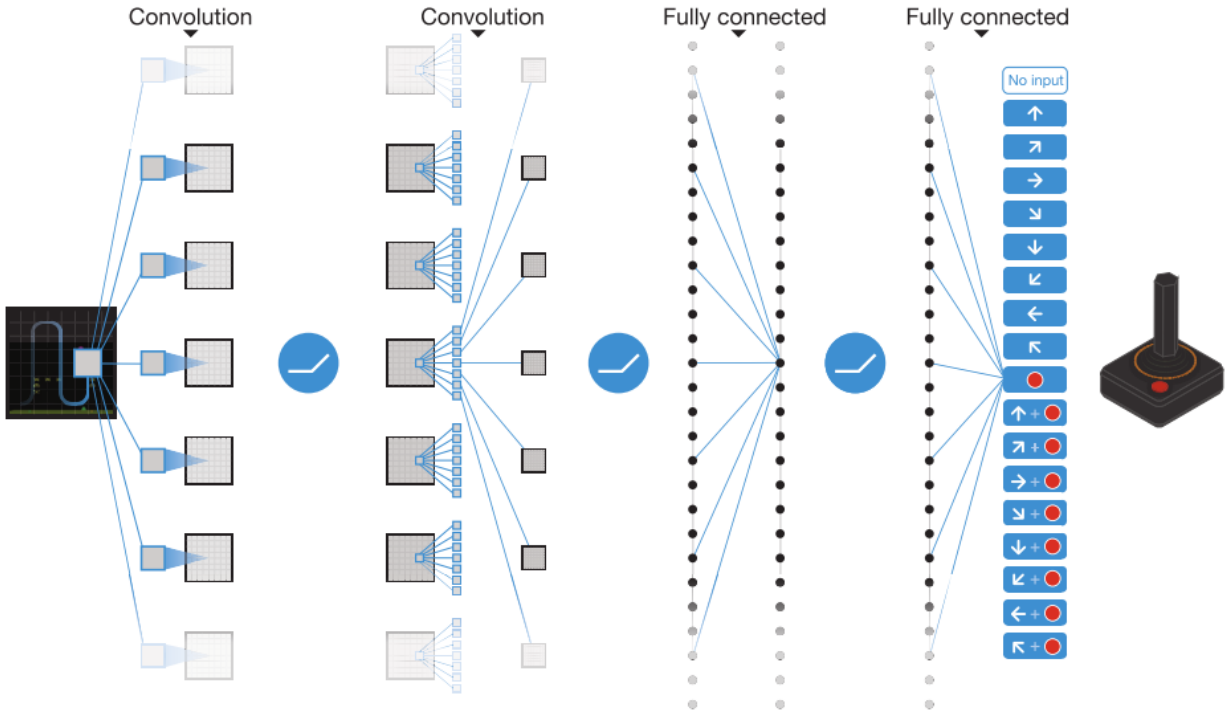
Deep Reinforcement Learning on Atari Games



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

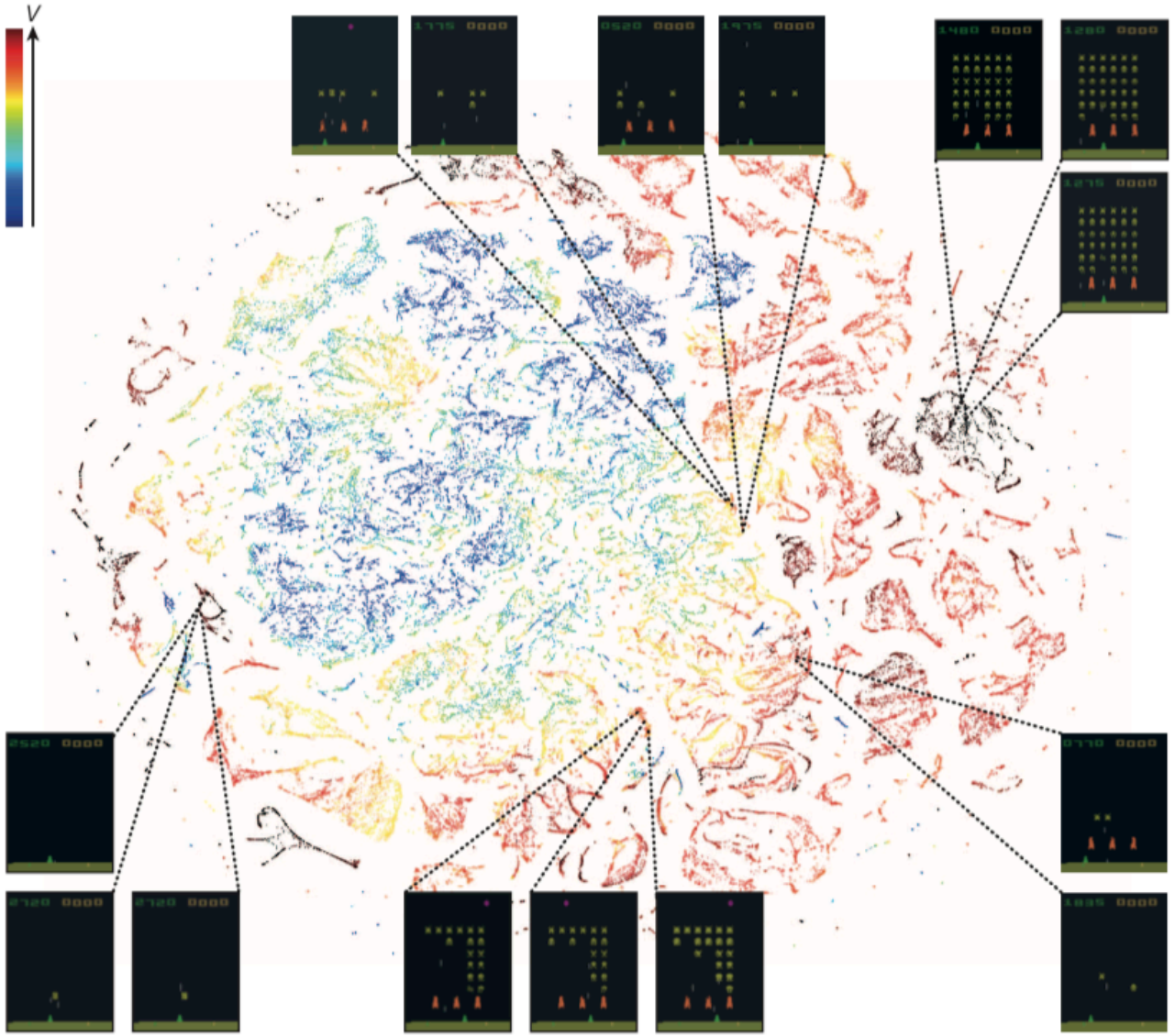
Many algorithms since then trying to provide a better way to learn the value function with DNNs

- Alpha Go – RL + human training
- Alpha Zero – RL + MCTS search + playing itself (Go, Chess)
- AlphaStar – RL + LSTMS + ? = play StarCraft against human experts



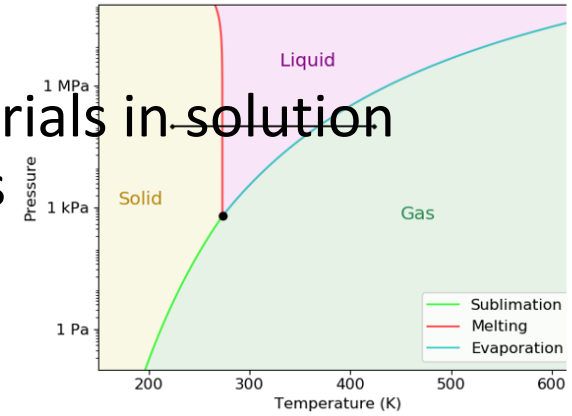
RL as Learning a Domain Generalization

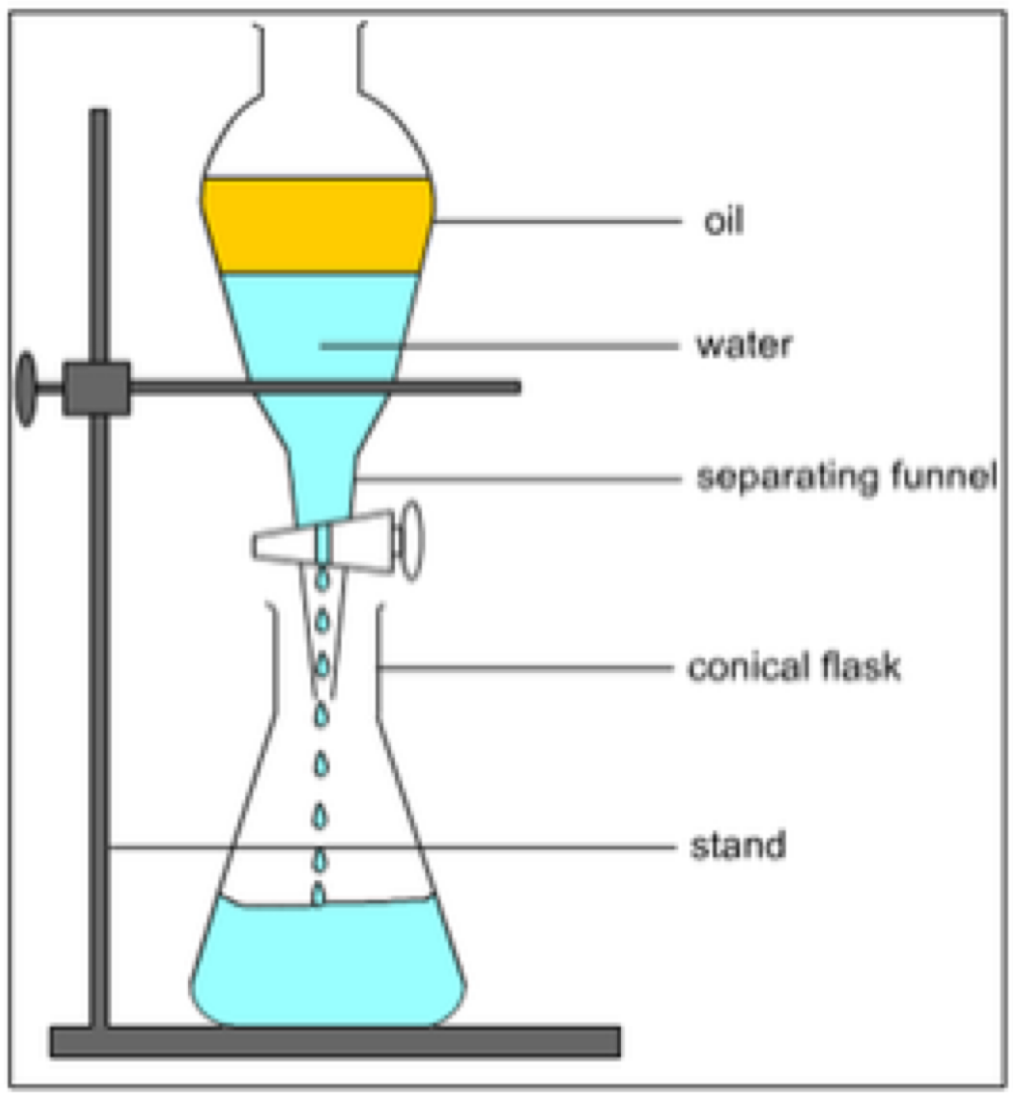
- Generalizing across “games” to learn to act robustly in new, similar situations
- Learning a causal model of a complex system by interacting with
- Bringing science and experimentation into AI
- What better way to do this than with actual physical science?

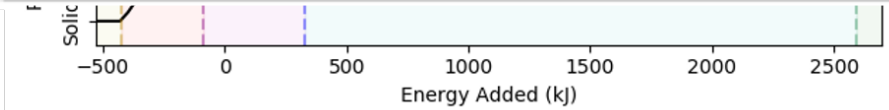


Multi-Stage Exploration

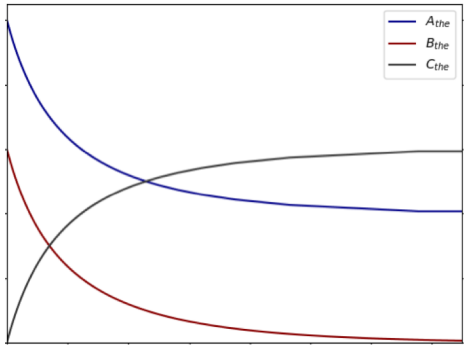
- We approach this by building small, manageable models for component activities in materials design and chemistry
  - (1) Reaction rates of various collections of chemicals at give temperatures and pressures
  - (2) Navigating the phases of matter of a given compound using temperature and pressure changes
  - (3) Extraction of materials in solution via polarizing solvents

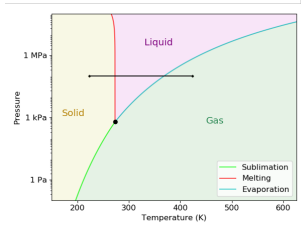


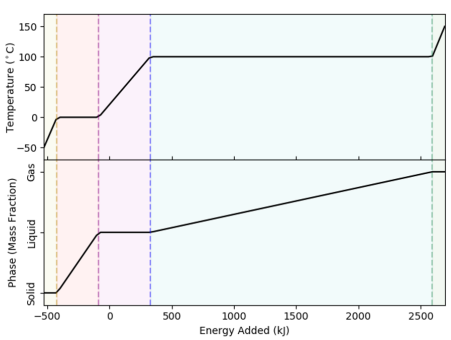


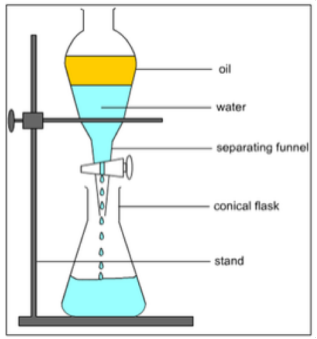


Multi-Stage Exploration









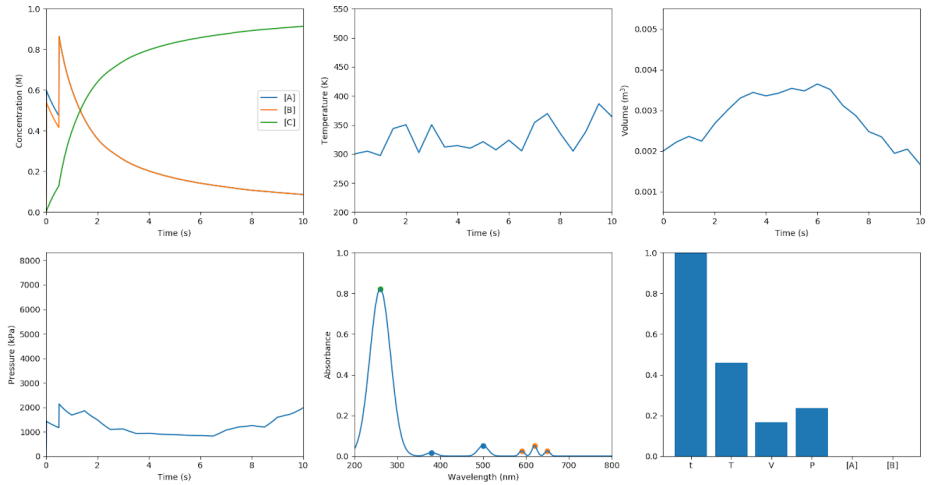
Automated Material Synthesis using Deep RL

WATERLOO.AI
WATERLOO ARTIFICIAL INTELLIGENCE INSTITUTE

Chemistry Lab

Process Actions

- (1) ODE-World
- (2) Phase-World
- (3) Extraction-World



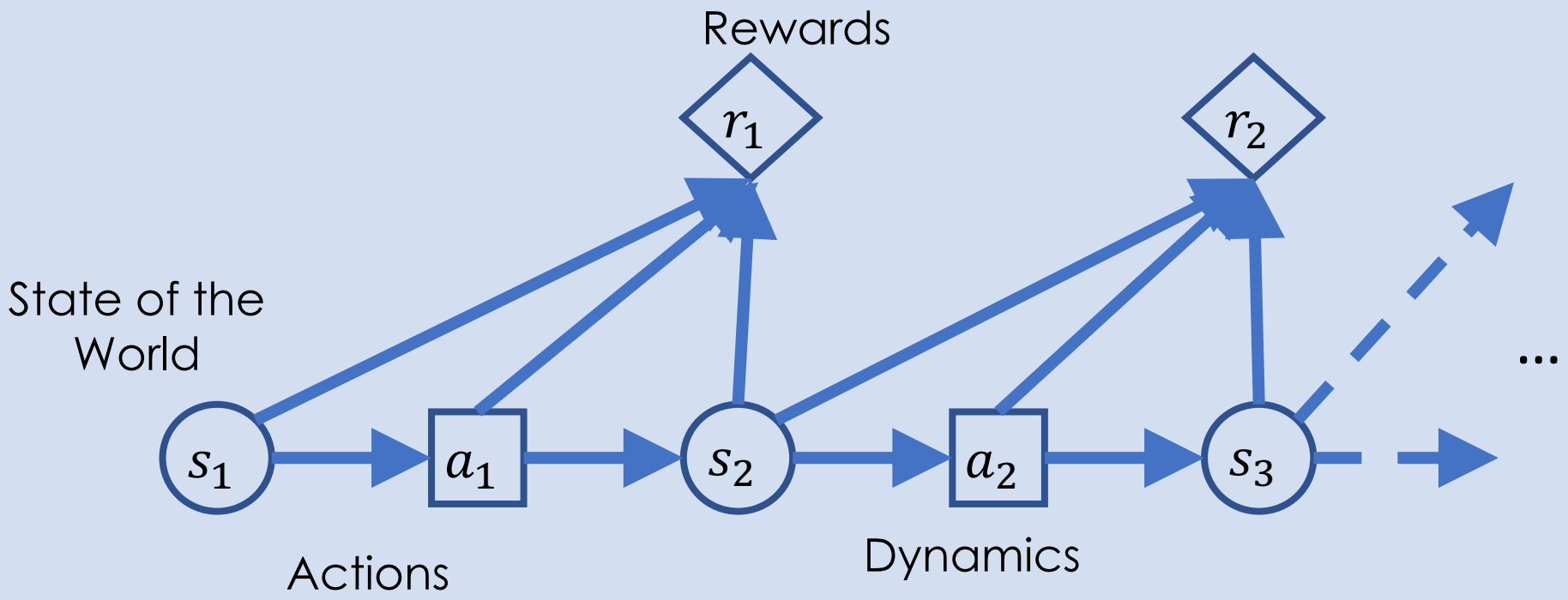
Observation Actions

- Mass Spectrometry*
- Nuclear Magnetic Resonance*
- Gas Chromatography*
- High Performance Liquid Chromatography*
- UV-Vis Spectrometry*
- Fluorescence Spectrometry*

Thank you – Questions and Discussion

Different Fields of AI Come Down to (one way to look at it)

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



Waterloo.AI — The Waterloo AI Institute

- Joint initiative of Faculty of Engineering and Faculty of Mathematics

UNIVERSITY OF WATERLOO
FACULTY OF ENGINEERING

UNIVERSITY OF WATERLOO
FACULTY OF MATHEMATICS

- Includes over 100 researchers in AI:
  - management sciences, electrical and computer, mechanical and mechatronics, systems design, civil and environmental, and chemical engineering
  - computer science, statistics and actuarial sciences, combinatorics and optimization
  - public health, health systems, biology, chemistry, earth and environmental sciences, and physics and astronomy, economics, accounting and finance

Waterloo.ai – Multidisciplinary Research

- Multidisciplinary research teams including mathematicians, computer scientists, and engineers
- Established expertise in collaborating with industry and developing real-world solutions to commercial challenges

WATERLOO.AI
WATERLOO ARTIFICIAL INTELLIGENCE INSTITUTE

Waterloo.ai — Foundational Research Areas

- **Machine learning, statistical learning**
- Natural language processing
- **Computer vision**
- **Probabilistic models, knowledge discovery, and knowledge representation**
- **Multi-agent systems** and game theory
- Health Informatics
- **Optimization and decision making**
- Trust modeling
- Affective computing and sentiment analysis
- Human-computer interaction
- Neuroscience