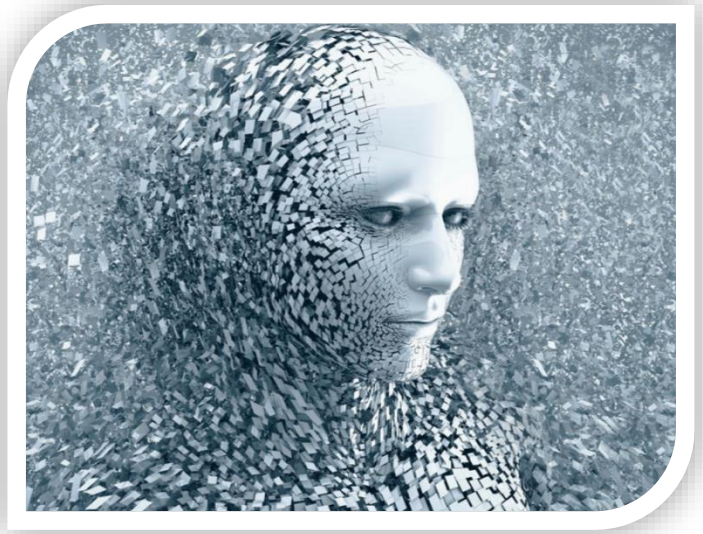


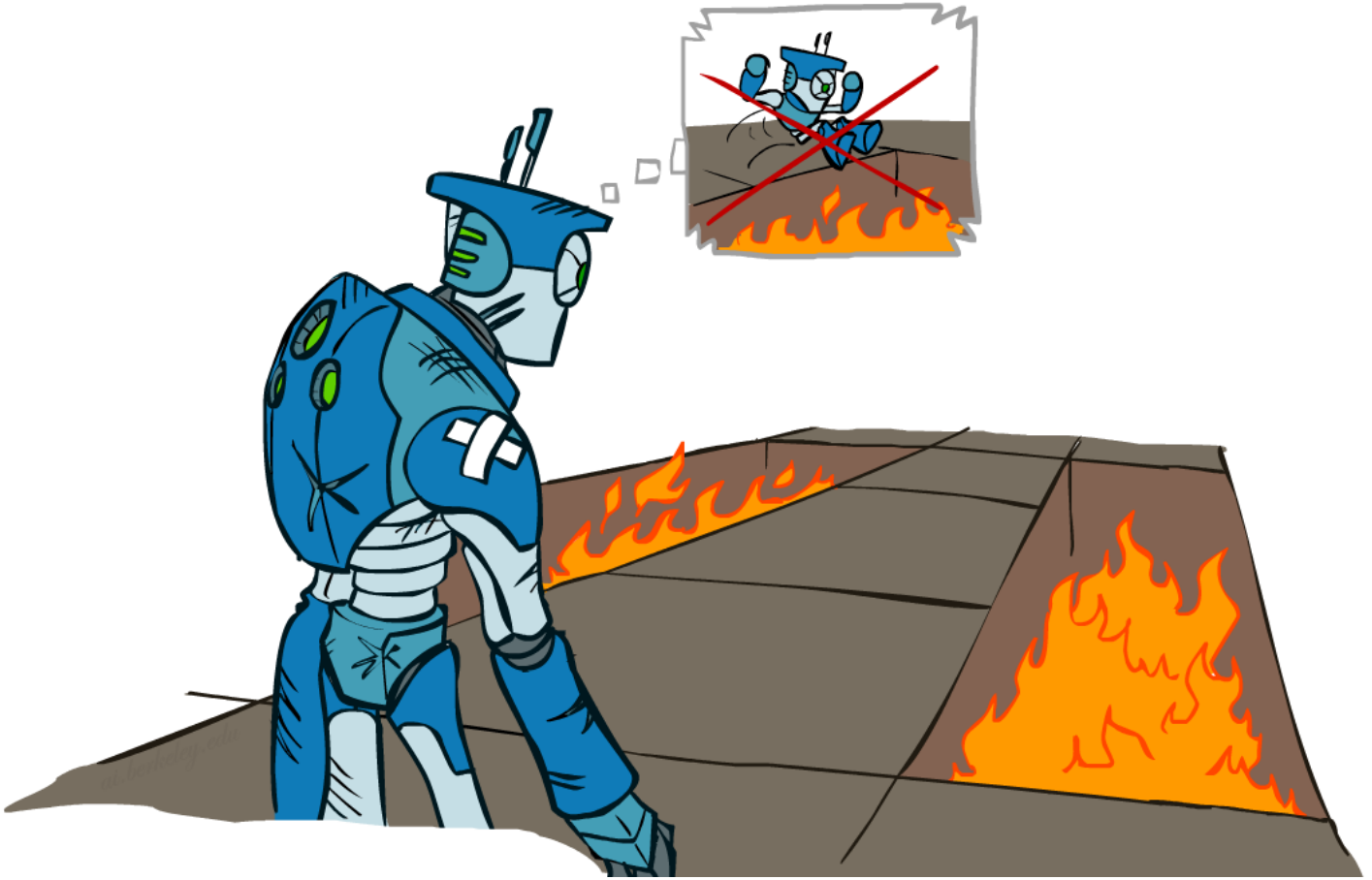
Artificial Intelligence

Session 6:
Learning

School of Computing and Engineering
University of West London, UK

Dr Massoud Zolgharni

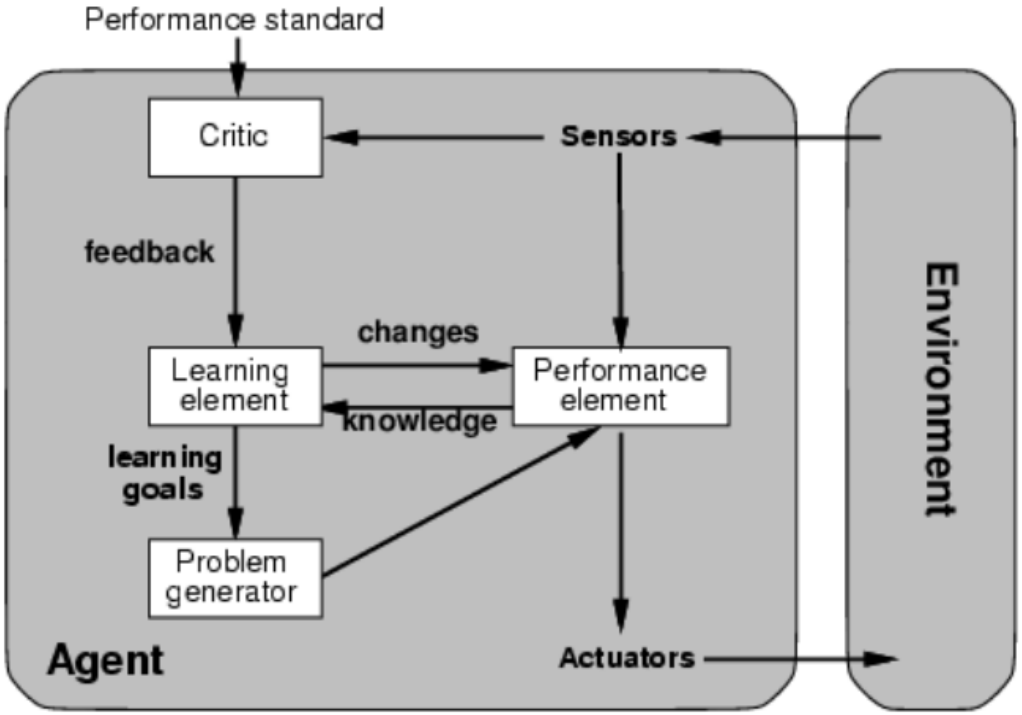
Learning



What is learning?

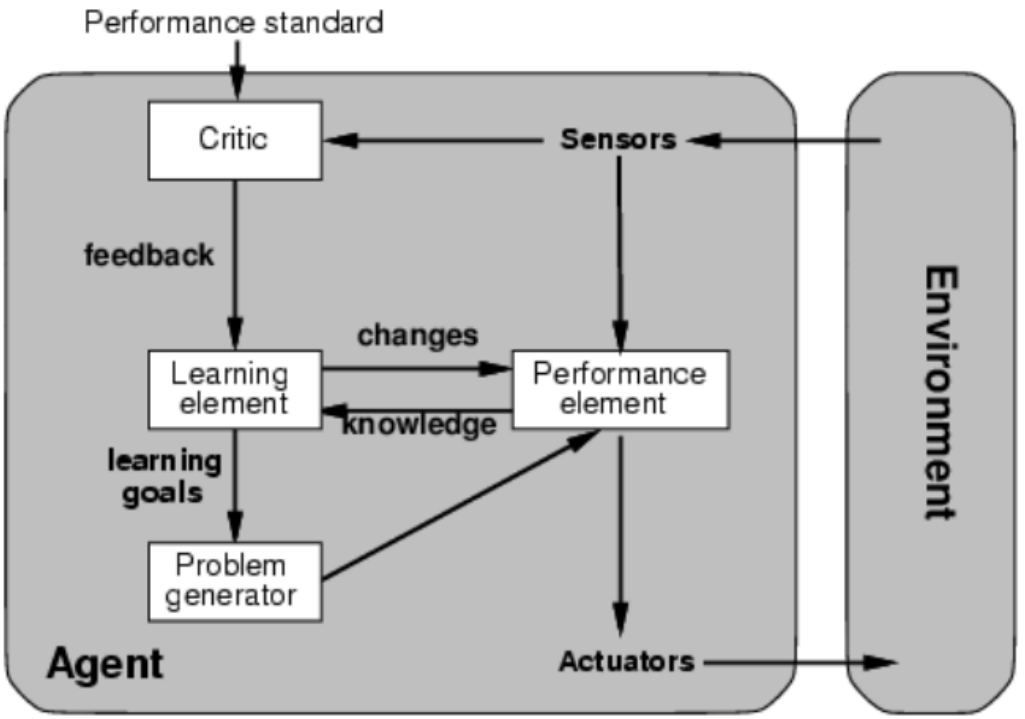
In intelligent agents, percepts should be used not only for acting, but also for improving the agent's ability to act in the future.

A learning agent has a performance element and learning element.



Learning allows an agent to **adapt to new circumstances** and to **detect and extrapolate patterns**.

What is learning?



Learning Element:

Adds knowledge, makes improvement to system

Performance Element:

Performs task, selects external actions

Critic:

Monitors results of performance, provides feedback to learning element

Problem Generator:

Actively suggests experiments, generates examples to test

Performance Standard:

Method / standard of measuring performance

Forms of learning

Learning depends on four major factors:

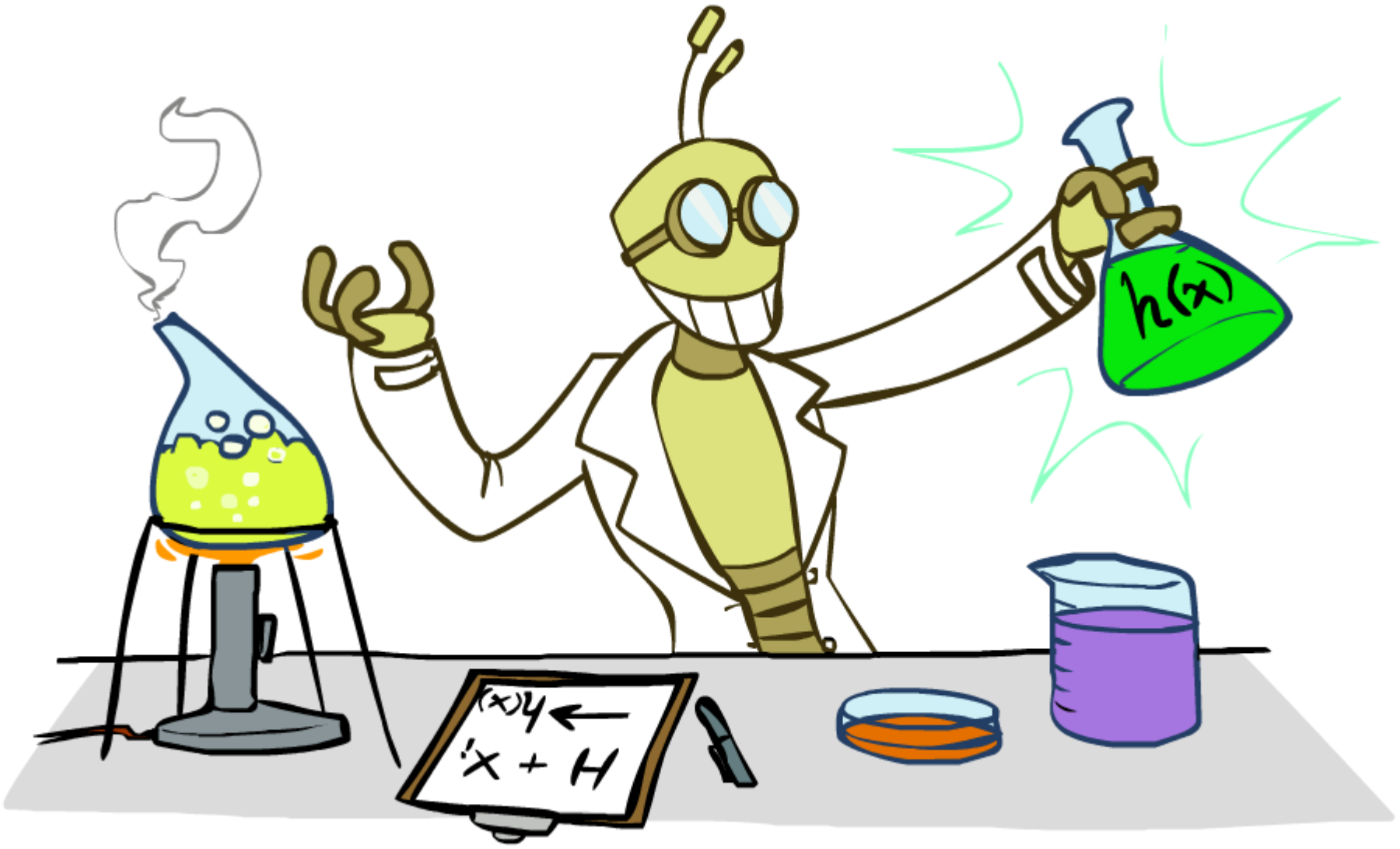
- which component(s) in performance element are to be improved
- what prior knowledge the agent already has (i.e., are we building the agent from scratch, or just updating it?
- the representation of the percept, e.g., numeric, character strings, sets of objects
- what feedback is available to learn from

Three main forms of learning in intelligent agents:

- ❑ Supervised learning (learning by examples)
- ❑ Unsupervised learning (learning by experimentation)
- ❑ Reinforcement learning (learning by being told; rewarded/punished)

The form of learning to use depends on the type of feedback

Supervised (Inductive) Learning



Supervised Learning

Given a **training set** of N example **input-output pairs**

$$(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N),$$

where each y_i was generated by an unknown function $y = f(x)$, **discover a function** h that approximates f.

The goal is to **predict the output value** y^* for a new unseen before input x^*.

The availability of a **target output value** y_i for each input x_i differentiates supervised learning from other forms of learning.

The function h is called a **hypothesis**. Supervised learning can be viewed as searching in the space of hypotheses for the best h.

Supervised Learning

Supervised learning: learning from examples

Highly simplified model

Examples are pairs of (x, $f(x)$)

f is the target function

Problem: find a hypothesis h

Such that $h \approx f$
Given a training set of examples

$x = 1\ ,\ f(x) = 1.2$

$x = 2\ ,\ f(x) = 2.4$

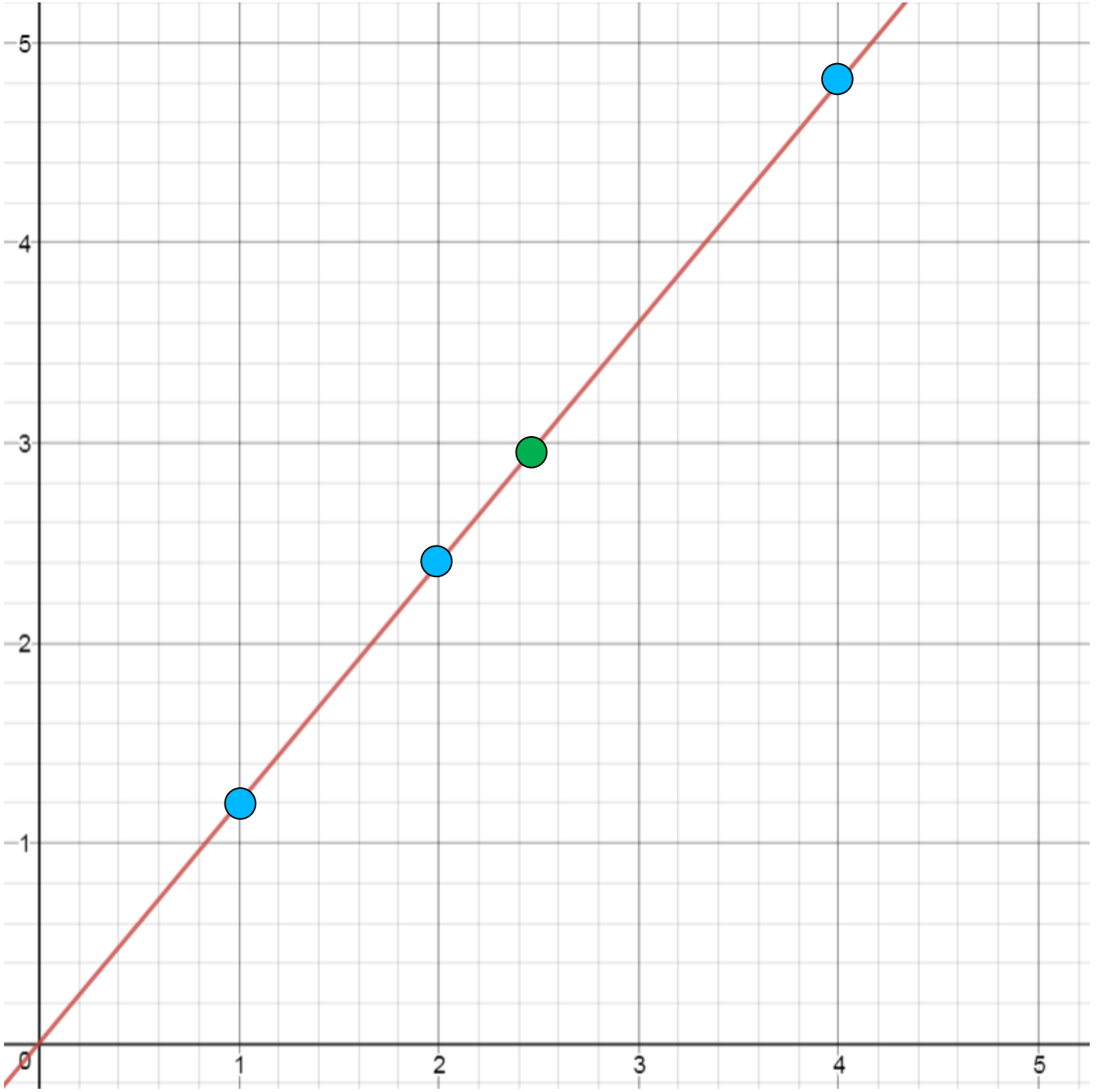
$x = 4\ ,\ f(x) = 4.8$

…

$f = ?$

$f(x) = 1.2 \times x$

Supervised Learning



$x = 1$, $f(x) = 1.2$

$x = 2$, $f(x) = 2.4$

$x = 4$, $f(x) = 4.8$

…

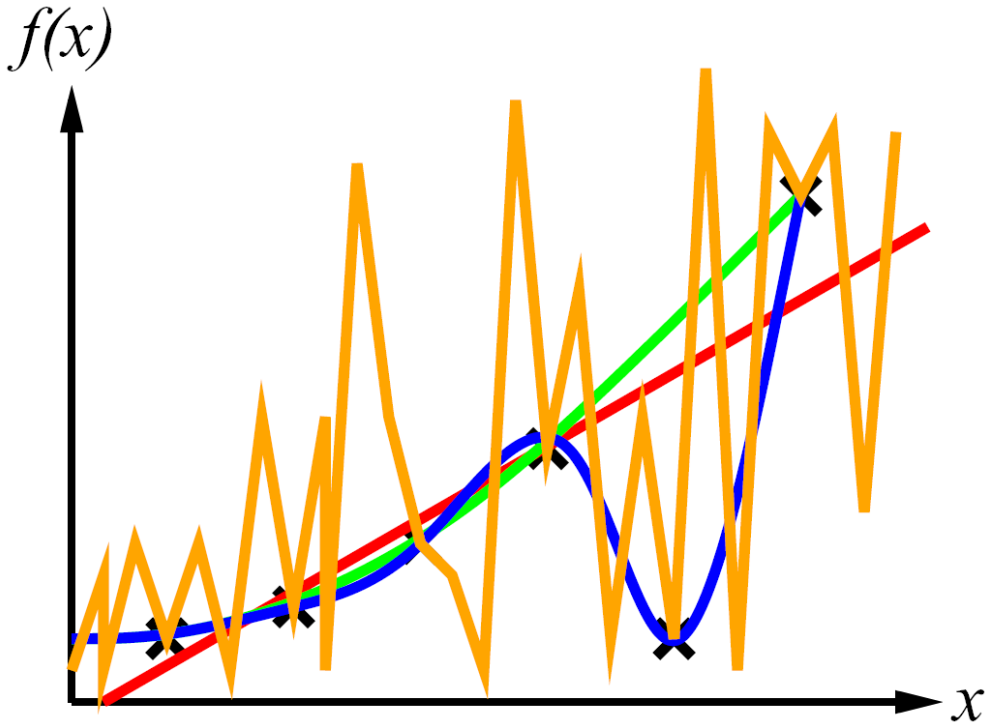
$f = ?$

$f(x) = 1.2\ x$

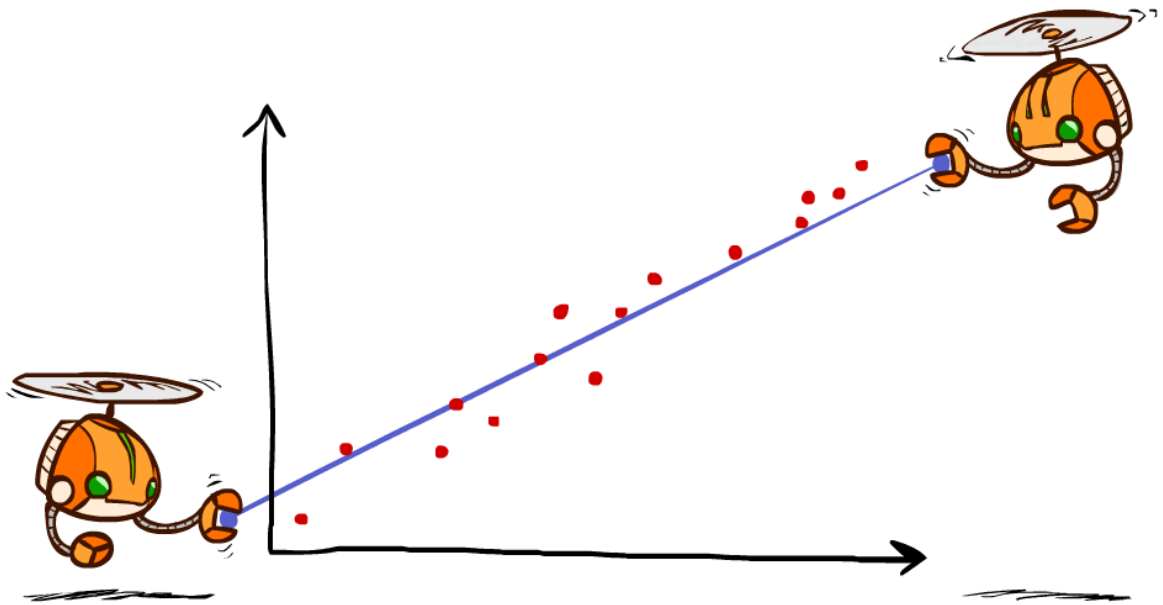
$x = 2.5$, $f(x) = 3$

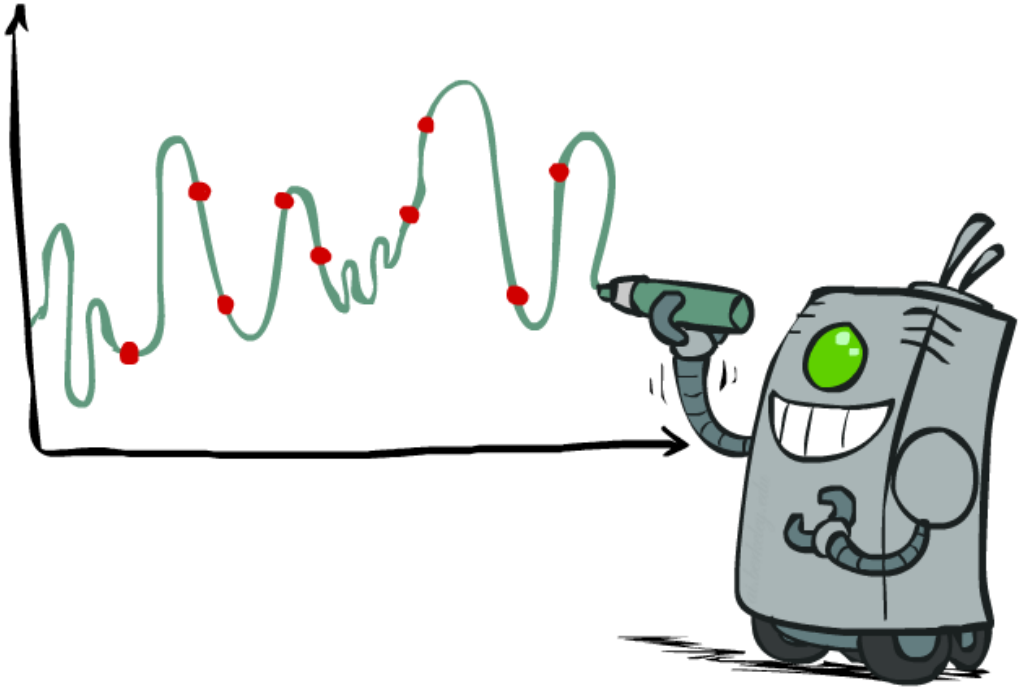
Supervised Learning

- Curve fitting (regression, function approximation):



Supervised Learning



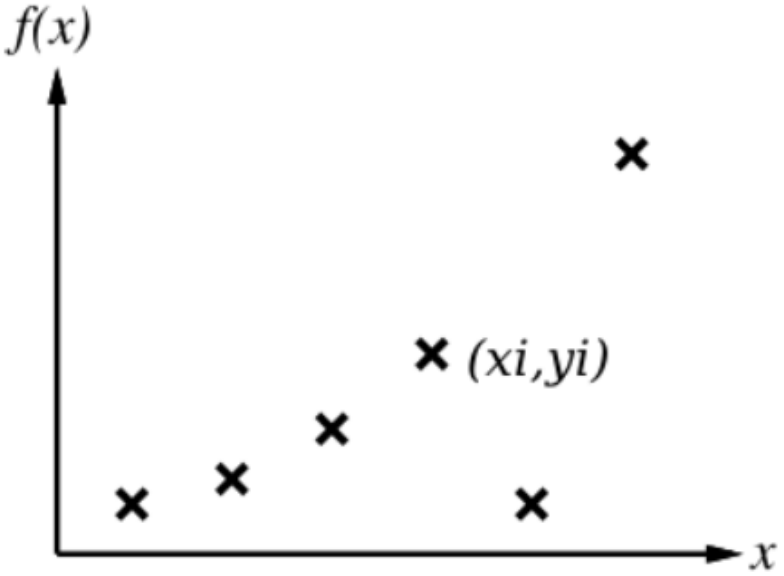


- Consistency vs. simplicity
- Ockham's razor: prefer simplest hypothesis consistent with data

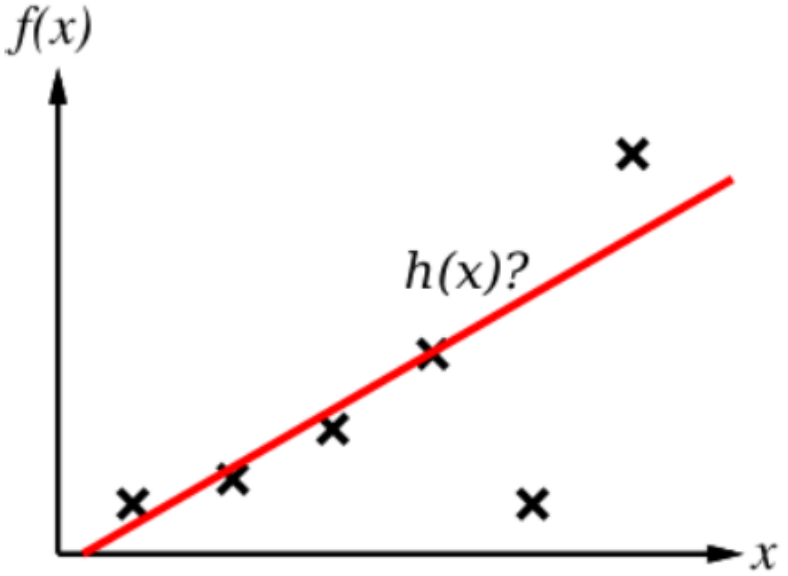
Regression Problem

The values for x_i, y_i can be numeric, discrete, continuous, non-numeric (nominal, symbols, text), etc.

When the target outputs are **continuous** values, we have a **regression** problem.



(a) Training set of input-output pairs. $f(x)$ is unknown.

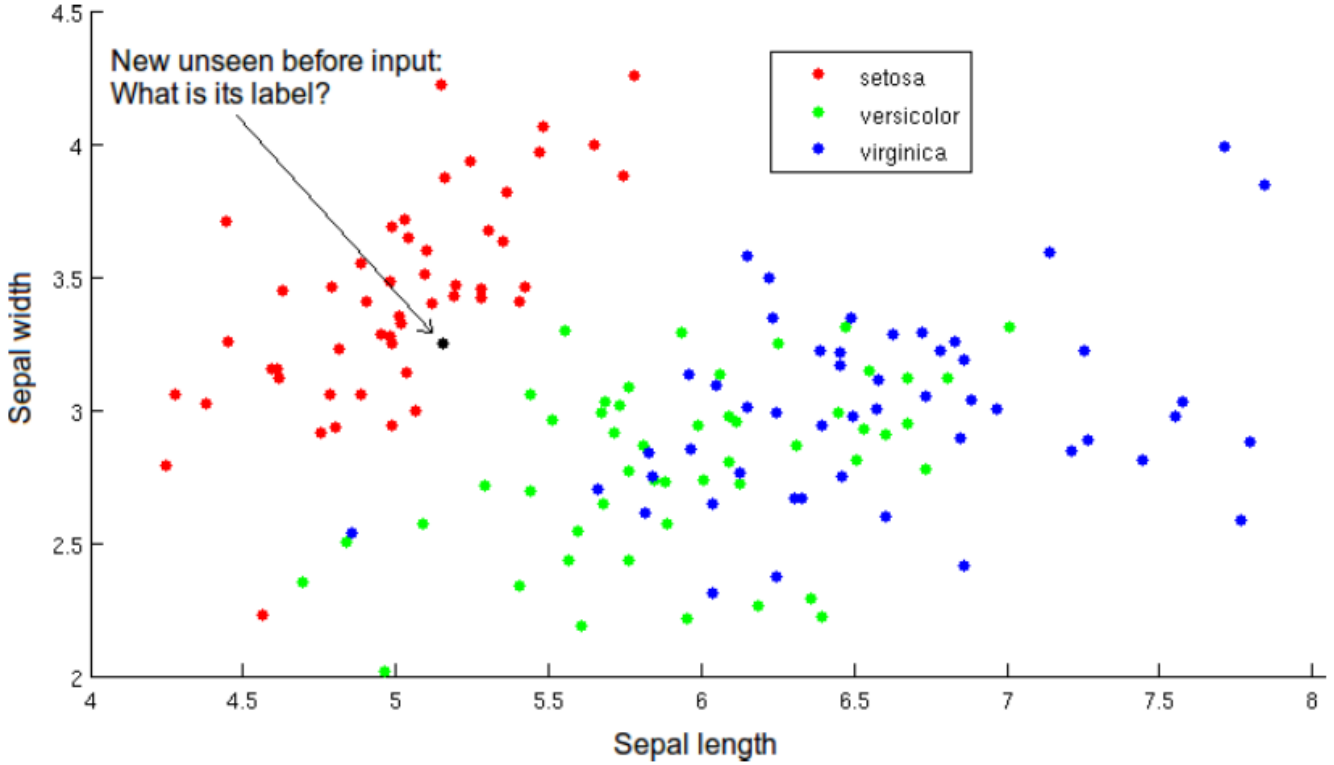


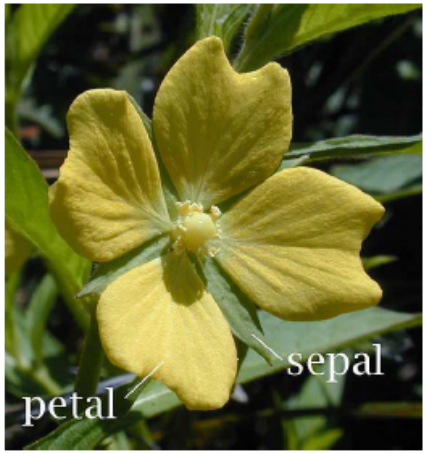
(b) Candidate hypothesis $h(x)$.

Classification Problem

When the target outputs are nominal values or **class labels**, we have a **classification** problem.

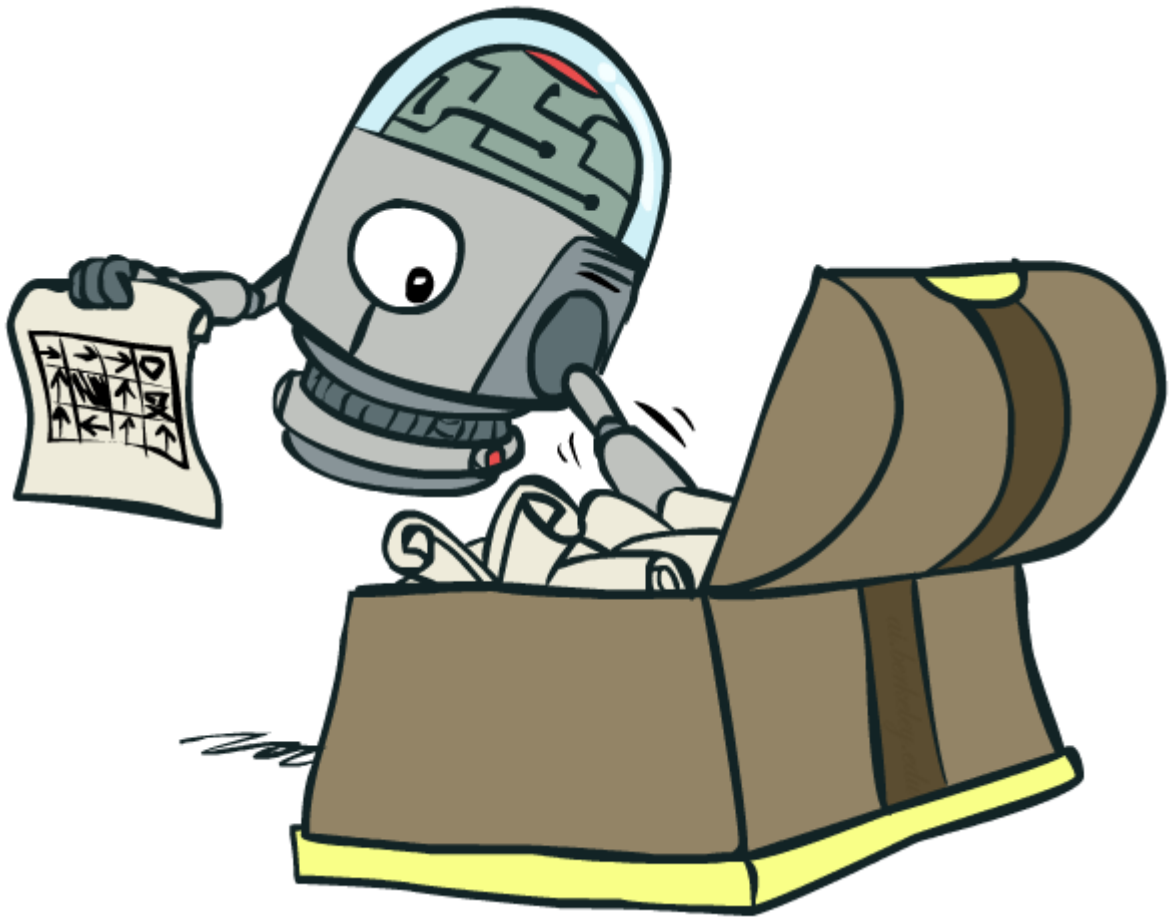
Example: Given measurements of sepal length and sepal width, identify the **type** of flower.





Here each input value $\mathbf{x}_i$ is a two-dimensional vector, while each target value y_i can be setosa, versicolor or virginica.

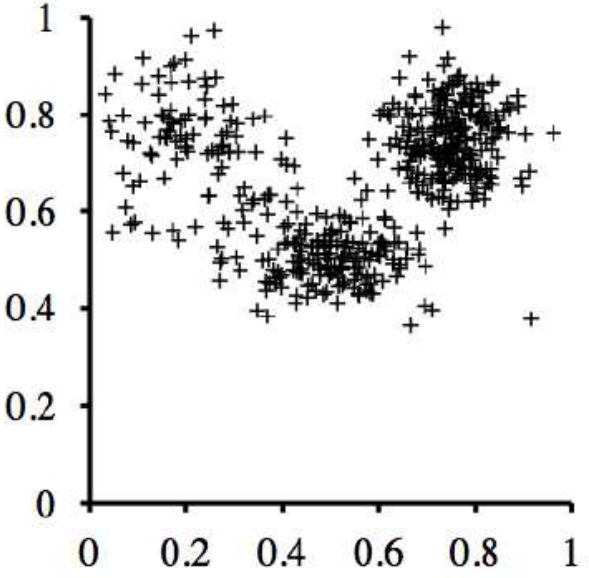
Unsupervised Learning



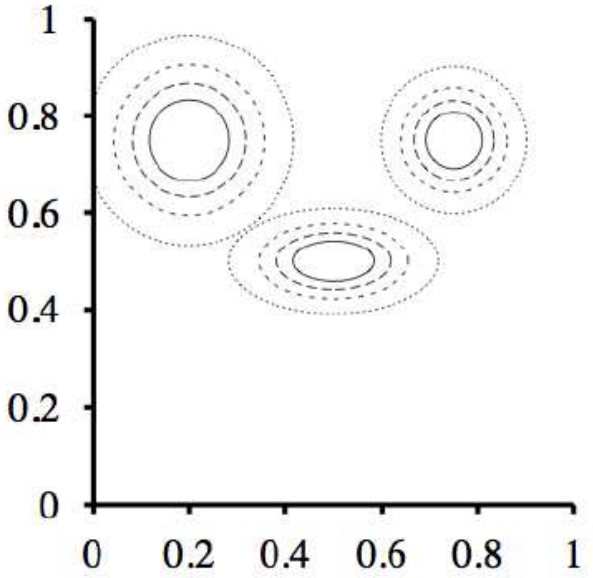
Unsupervised Learning

Learning **patterns** in the input when no specific target output values are supplied.

Example: Finding clusters in 2D input data.



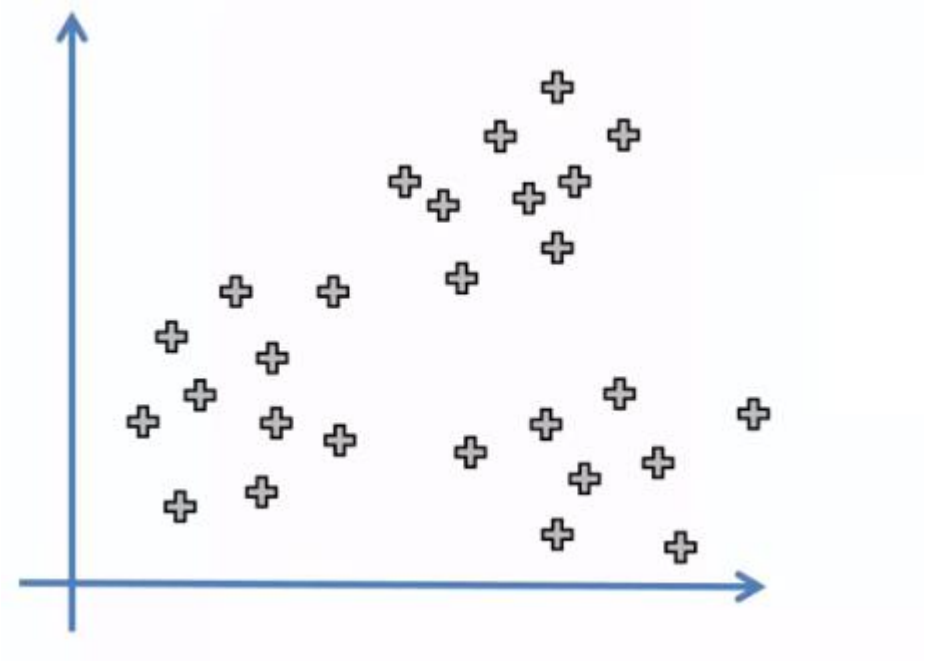
(a) Input data with unknown cluster memberships.



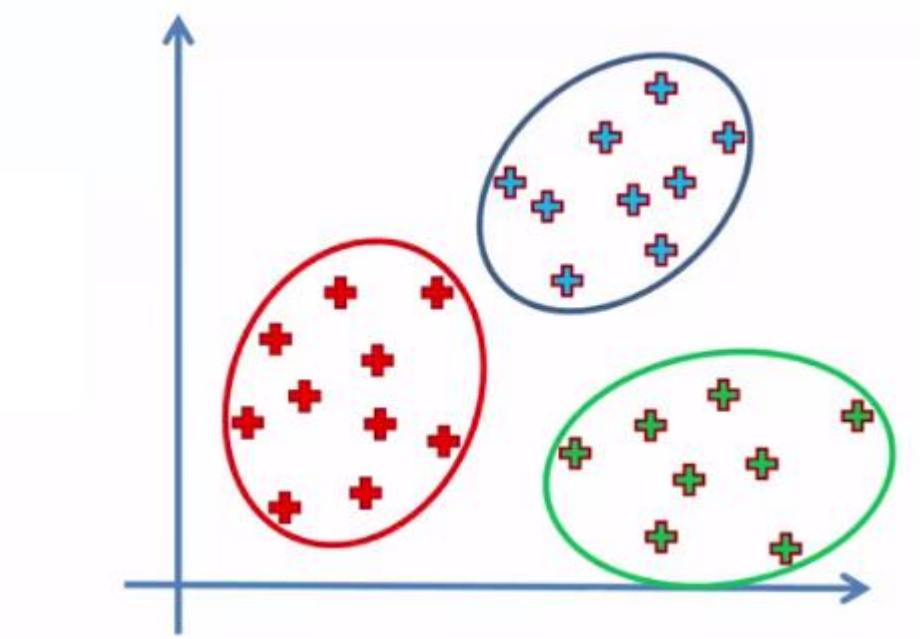
(b) A hypothesis clustering.

Unsupervised Learning

- We manufacture and sell trousers.
- We have data from our customers, and plotted waist size vs. height in the graphs below.
- While there could be all kinds of combinations, we can't manufacture every possible size. However, we know we could target most of our customers by making a few specific sizes.
- Cluster the data, observe where the clusters are cantered, and just manufacture trousers in those sizes.
- That way, we can appeal to most of our customers while using our manufacturing resources most effective.



Raw data

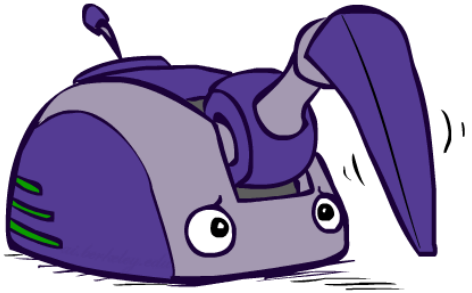


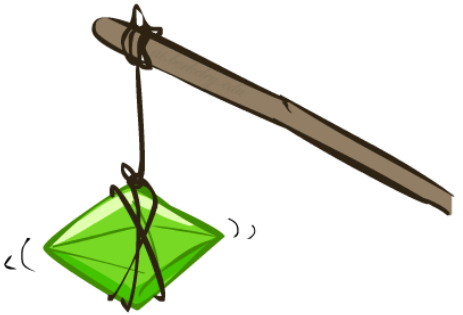
Clustered data

Reinforcement Learning

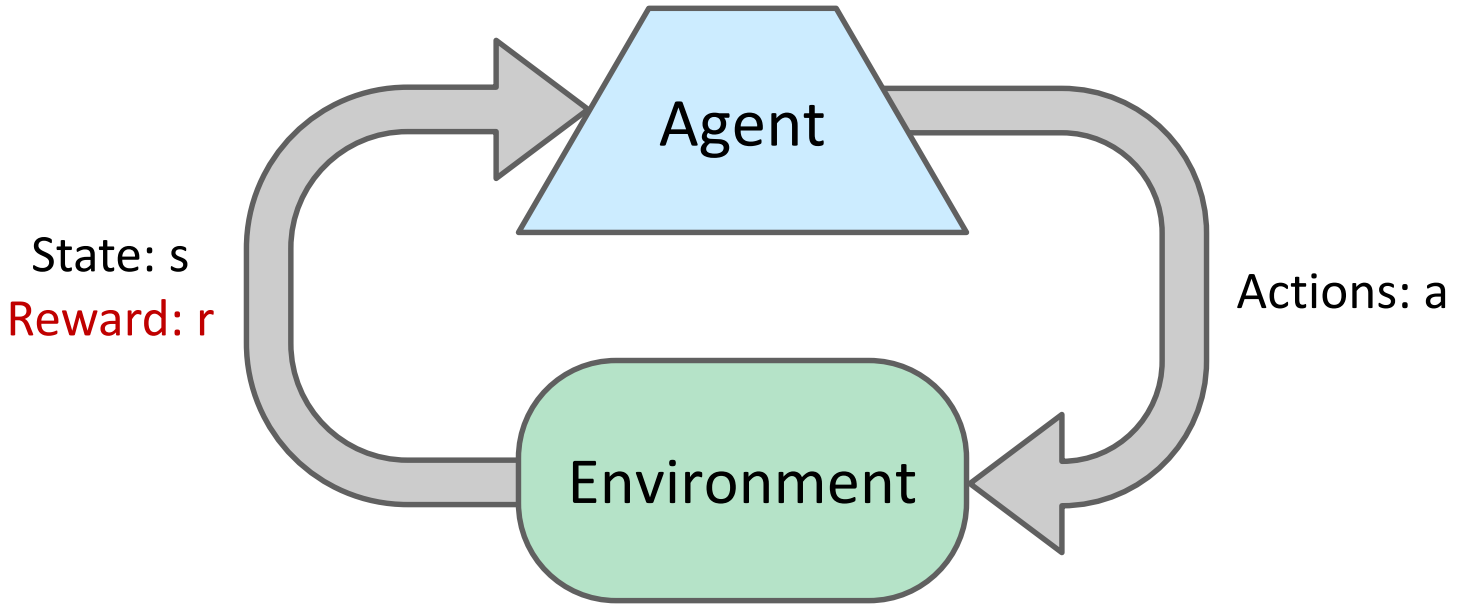
Reinforcement Learning

Learning behaviour (optimise behaviour for rewards)





Reinforcement Learning



- Basic idea:
  - Agent performs an action
  - Receive feedback in the form of rewards
  - Agent's utility is defined by the reward function
  - Must (learn to) act so as to maximize expected rewards (utility)
  - All learning is based on observed samples of outcomes!

Learning to Walk

Initial

A Learning Trial

After Learning [1K Trials]

[Kohl and Stone, ICRA 2004]

Learning to Walk

Initial

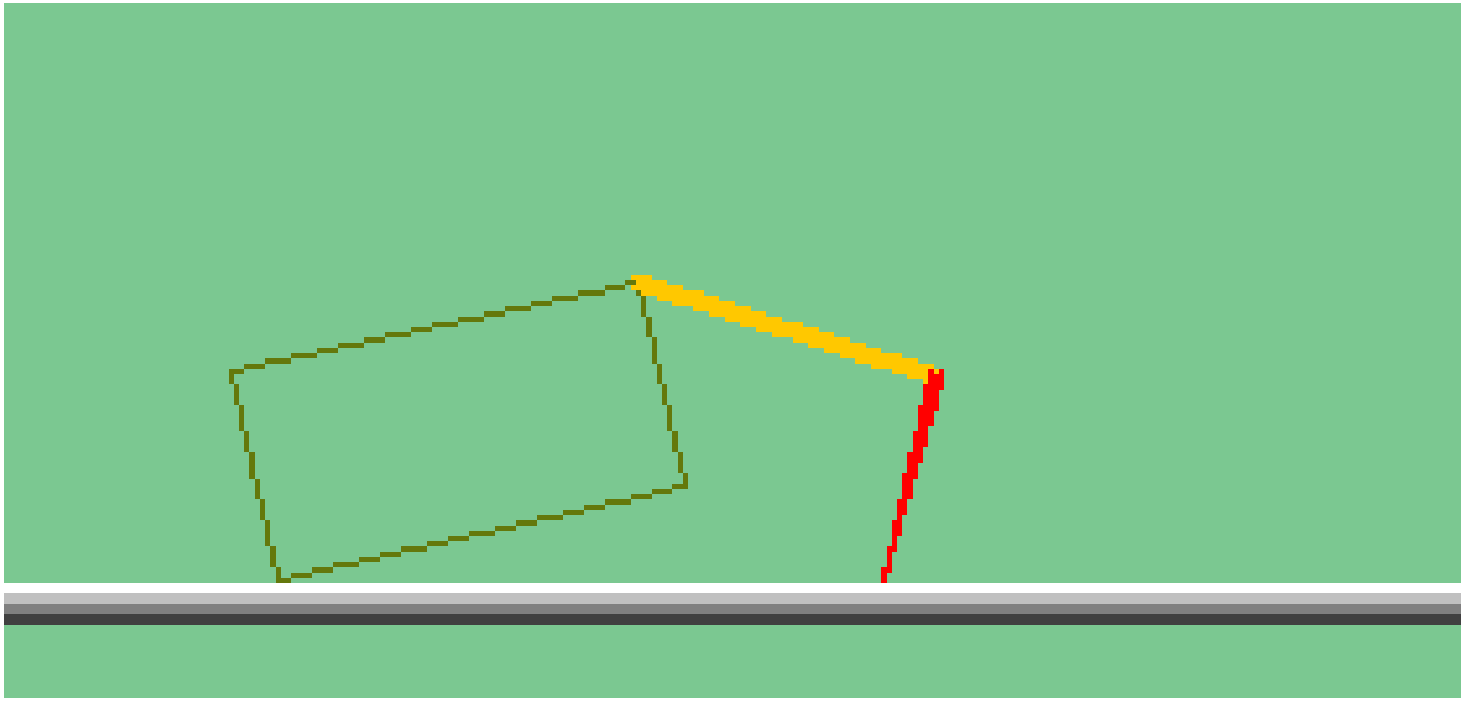
Learning to Walk

Training

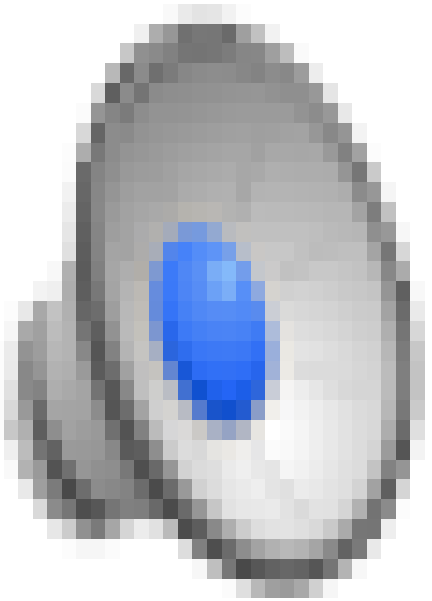
Learning to Walk

Finished

The crawler!

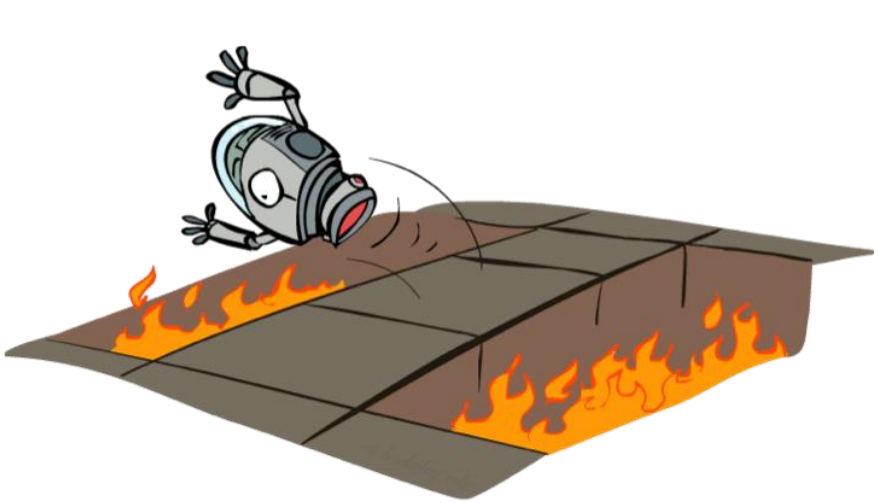


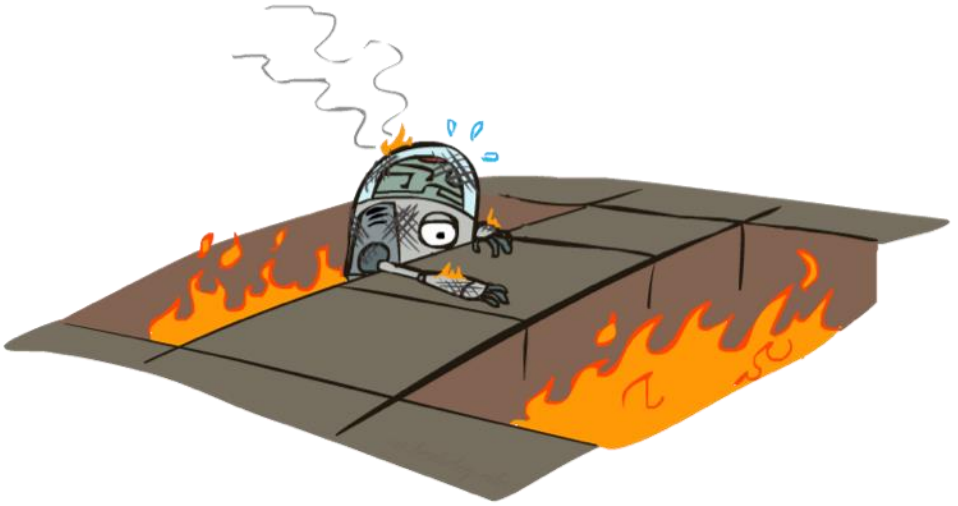
The crawler!



Example: Toddler Robot

Reinforcement Learning





- The positive or penalty (negative reward) is given after a sequence of actions
- For example, the agent does not know the fire is bad until it has taken action and experienced it
- A lot of times, we do simulations instead (cost of penalty could be high)

Summary

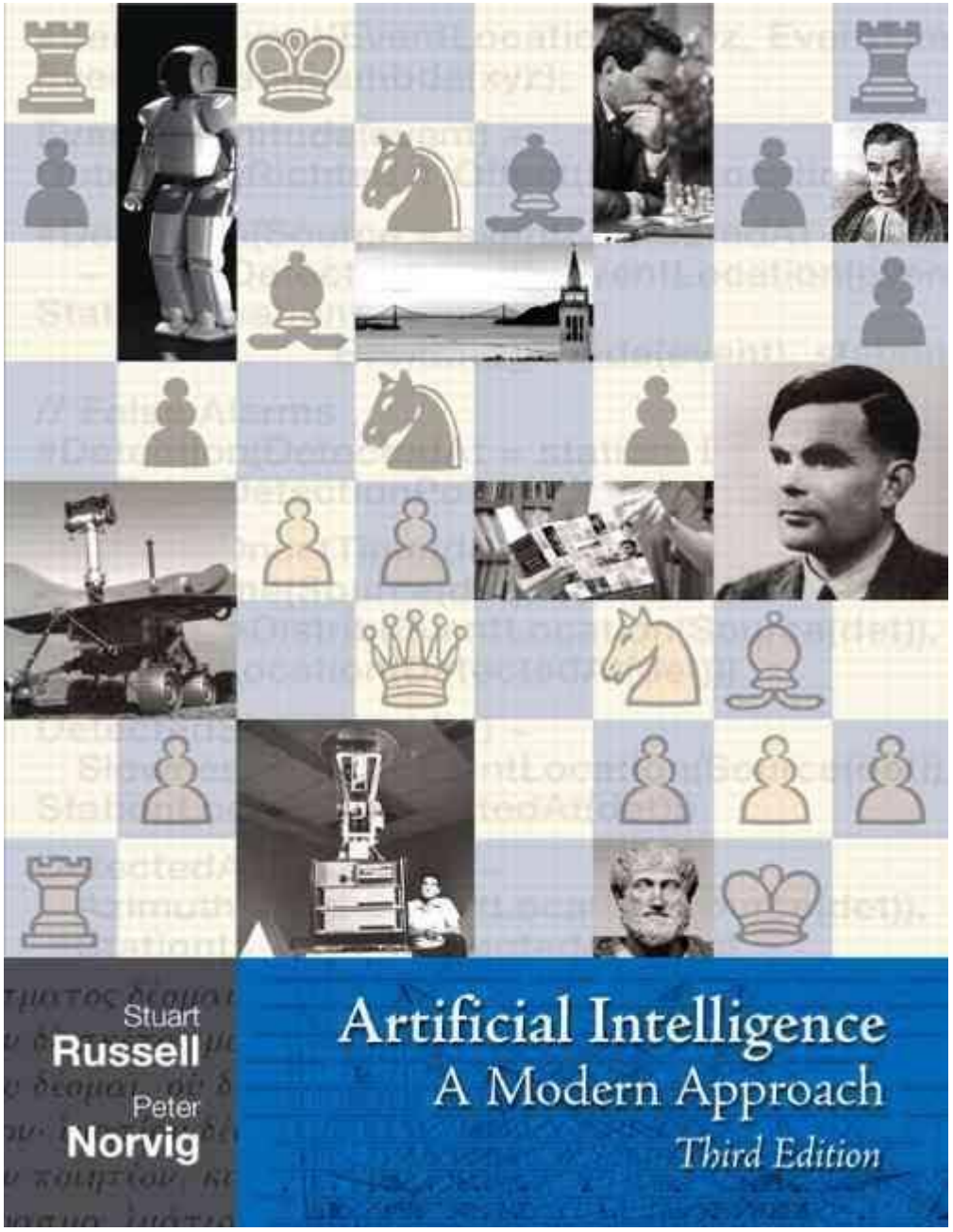
Three learning approaches:

- ✓ **Unsupervised learning**: an unsupervised model provides unlabelled data that the algorithm tries to make sense of by extracting features and patterns on its own
- ✓ **Supervised learning**: in a supervised learning model, the algorithm learns on a labelled dataset, providing an answer key that the algorithm can use to evaluate its accuracy on training data
- ✓ **Reinforcement learning**: reinforcement learning trains an algorithm with a reward system, providing feedback when an artificial intelligence agent performs the best action in a particular situation

Recommended reading

Stuart Russell, Peter Norvig: ***Artificial Intelligence A Modern Approach***

Chapter 18



ANY Questions?