Deep Learning for Robotic Control

Wei Qi Yan

Auckland University of Technology

Table of Contents

Machine Intelligence

- **Logic**: First-order logic, fuzzy logic, predicative logic, propositional logic, etc.
- **Search**: Uninformed search, informed (heuristic) search, adversarial search, etc.
- **Reasoning**: First-order logic, forward/ backward chaining, probabilistic reasoning, Bayes' rule, dynamic Bayesian networks, etc.
- **Making decisions**: Decision tree, decision networks, expert system, sequential decision, game theory, etc.
- **Reinforcement learning**: Passive/active RL, Q-learning, Markov decision process (MDP), policy search, etc.
- **Neural networks**: Single /multilayer FFNN, etc.

P. Norvig, S. Russell (2016) Artificial Intelligence: A Modern Approach (3rd Edition), Prentice Hall.

Propositional Logic: Truth Table

- Logic true: 't'
- Logic false: 'f'
- Negation: '$\neg$' (not)
- Conjunction: '$\wedge$' (and)
- Disjunction: '$\vee$' (or)
- Implication: '$\Rightarrow$'
- Equivalence: '$\Leftrightarrow$'

A	B	(A)	$\neg A$	$A \wedge B$	$A \vee B$	$A \Rightarrow B$	$A \Leftrightarrow B$
t	t	t	f	t	t	t	t
t	f	t	f	f	t	f	f
f	t	f	t	f	t	t	f
f	f	f	t	f	f	t	t

W. Ertel (2017) Introduction to Artificial Intelligence. Springer International Publishing.

Machine Intelligence

Nautre-Inspired Machine Intelligence

- Physics-based algorithms (PBAs)
- Chemistry-based algorithms (CBAs)
- Biology-based algorithms (BBAs)

Biology-Inspired Machine Intelligence

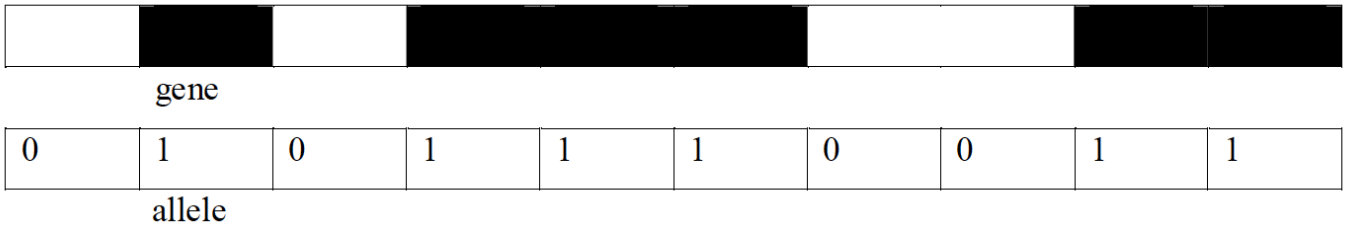
- Evolutionary Algorithms (EA)
- Brain-inspired Algorithms (BIA)
- Swarm Intelligence-based Algorithms (SIA)

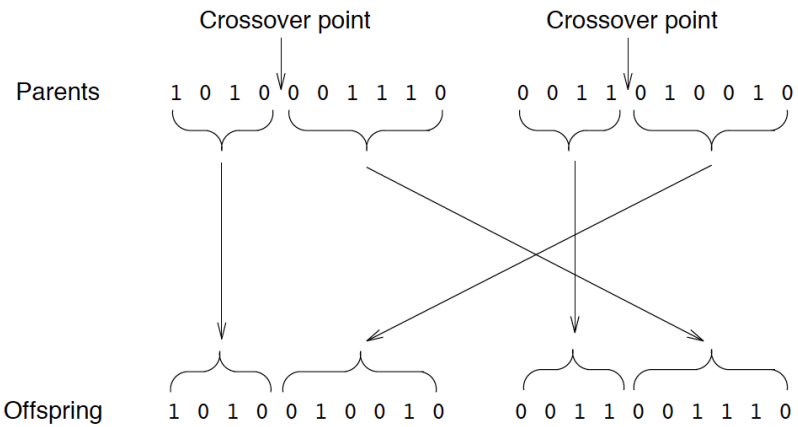
GA: Genetic Algorithm

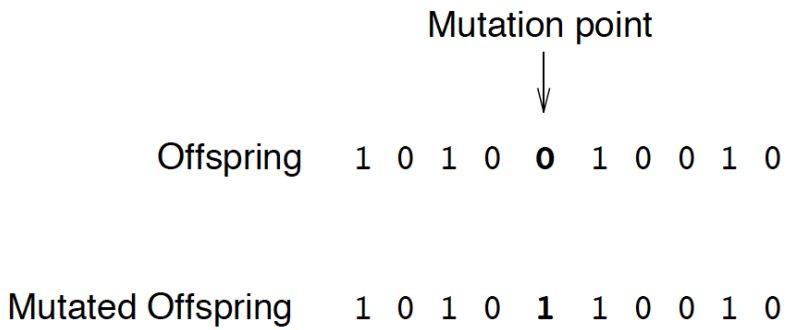
- **Initialization**: Create an initial population.
- **Evaluation**: Evaluate each member of the population, and calculate a "fitness" for the individual.
- **Selection**: Constantly improve populations' overall fitness.
- **Crossover**: Create new individuals by combining aspects of the selected individuals.
- **Mutation**: Add randomness into populations' genetics.
- **Repeat**: Start again until a termination condition is reached.

P. Norvig, S. Russell (2016) Artificial Intelligence: A Modern Approach (3rd Edition), Prentice Hall.

GA: Genetic Algorithm







GA Termination Conditions

- A solution is found that satisfies the minimum criteria.
- A fixed number of generations reached.
- An allocated budget (e.g., computational time, etc.) reached.
- The highest ranking solution's fitness has reached.
- A plateau no longer produces better results.
- Combinations of the above.

P. Norvig, S. Russell (2016) Artificial Intelligence: A Modern Approach (3rd Edition), Prentice Hall.

GA Algorithm

Algorithm 1. A version of the genetic algorithm (GA).

Input: original function, search space (the set of all individual solutions, or chromosomes), fitness function.

Optional input: good enough solution, from which the good enough fitness value is calculated.

1. **Create initial population** by randomly selecting a certain number of chromosomes from the search space;
2. **Evaluate** the fitness of the current population by calculating every chromosome's fitness value.
3. **While the termination criteria are not met, do:**

[Terminat. criteria: the best fitness value is good enough or the specified number of iterations is done, or ...]

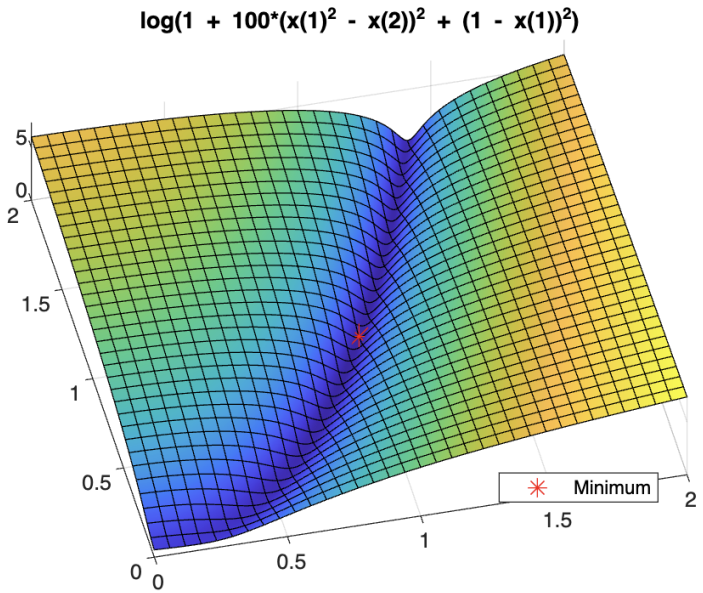
   - **3.1 While the new population is not completed, do:**
     - **3.1.1 Select** two parent chromosomes by random selection from the previous population;
     - **3.1.2 Recombine** the parent chromosomes to obtain two children;
     - **3.1.3 Enlist** the two children in the new population;
   - **3.2 Mutate** the chromosomes in the new population;
   - **3.3 Replace** the previous population with the new one;
   - **3.4 Evaluate** the fitness of the current population (0.2).

Output: the best chromosome found (the one with the highest fitness value).

P. Norvig, S. Russell (2016) Artificial Intelligence: A Modern Approach (3rd Edition), Prentice Hall.

MATLAB GA Algorithm

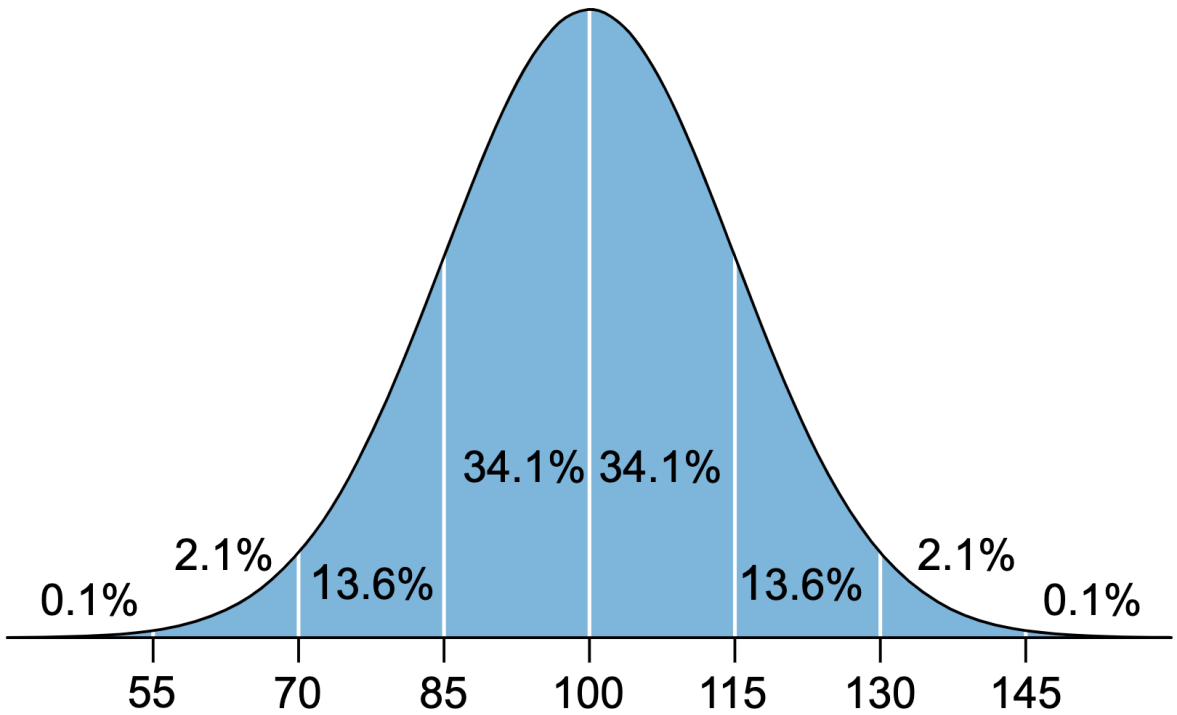
- With global optimum
- Without continuity
- Without derivatives
- Without linearity



https://au.mathworks.com/help/gads/fitness-function-forms.html

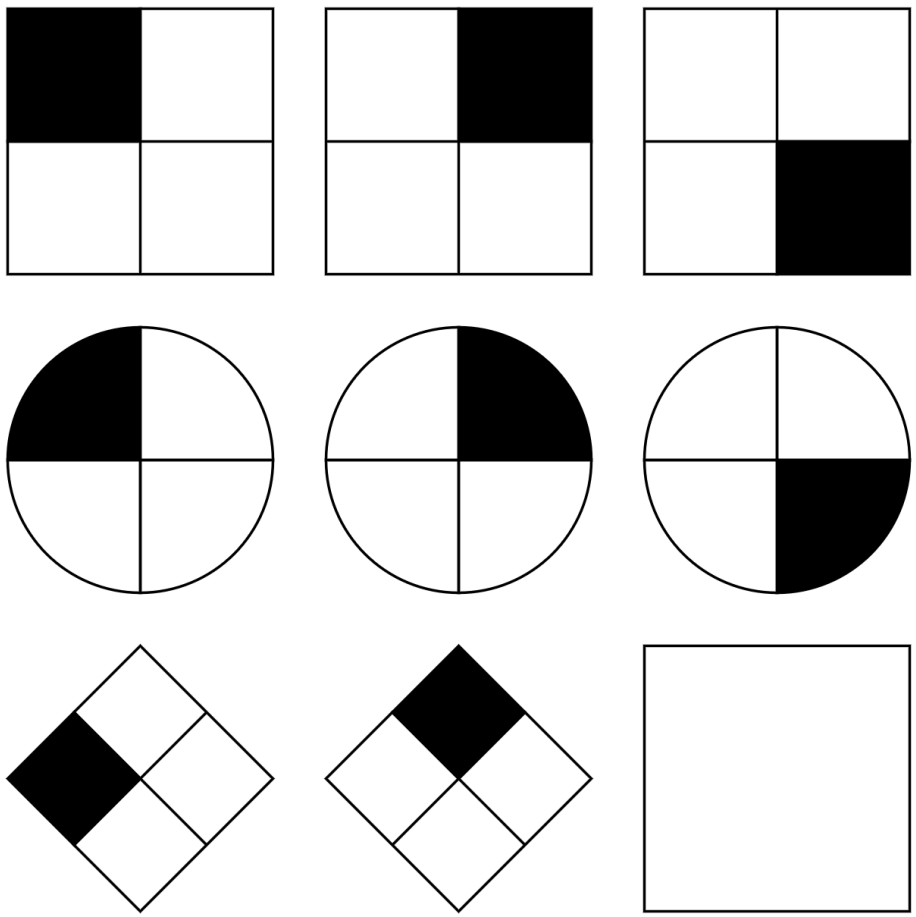
IQ: Intelligence Quotient

An intelligence quotient (IQ) is a total score derived from a set of standardised tests or subtests designed to assess human intelligence.



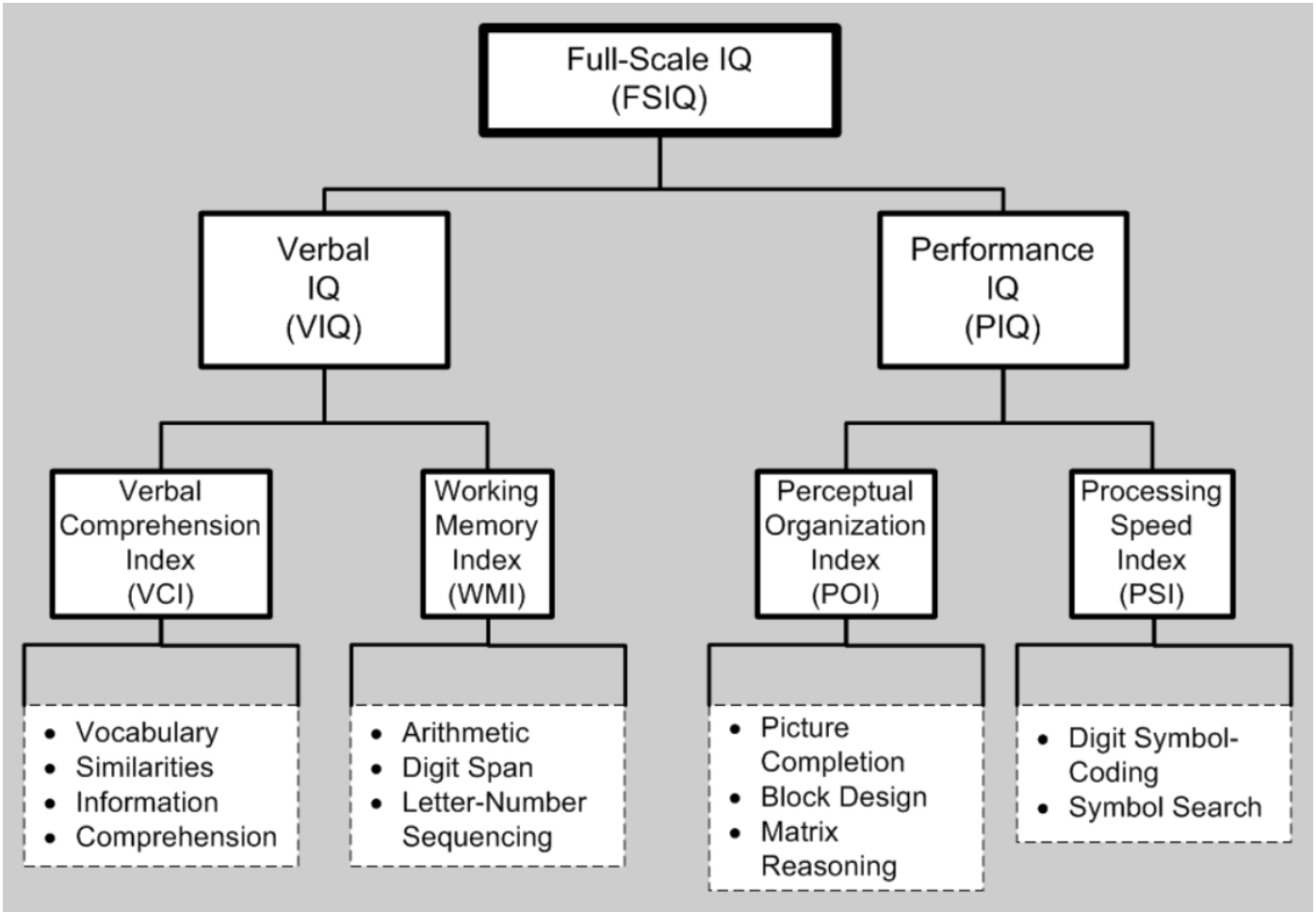
https://en.wikipedia.org/wiki/Intelligence_quotient

IQ: Raven's Progressive Matrices Test



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

WAIS-III: Wechsler Adult Intelligence Scale



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

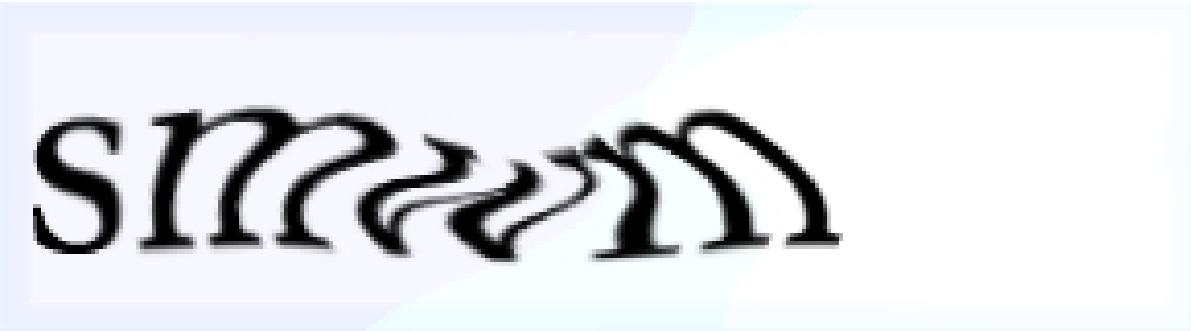
WAIS-5: Wechsler Adult Intelligence Scale (2024)

- The FSIQ is now generated from 7 subtests: Similarities, Vocabulary, Block Design, Matrix Reasoning, Figure Weights, Digit Span Sequencing, Coding.
- Fifteen ancillary index scores including the General Ability Index, are present.
- The test may be administered in the classic physical format or on a digital platform

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

Turing Test

- The Turing test is a test of a machine's ability to exhibit intelligent behaviour equivalent to, or indistinguishable from, that of a human.
- CAPTCHA stands for "Completely Automated Public Turing test to tell Computers and Humans Apart".
- ChatGPT-4 passes a rigorous Turing test, diverging from average human behavior chiefly to be more cooperative.



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

Questions?

Questions?

In GA algorithm, the termination conditions are:

1. A solution is found that satisfies the minimum criteria.
2. A fixed number of generations reached.
3. An allocated budget (e.g., computational time, etc.) reached.
4. None of the given options.

which answer is wrong?___

Questions?

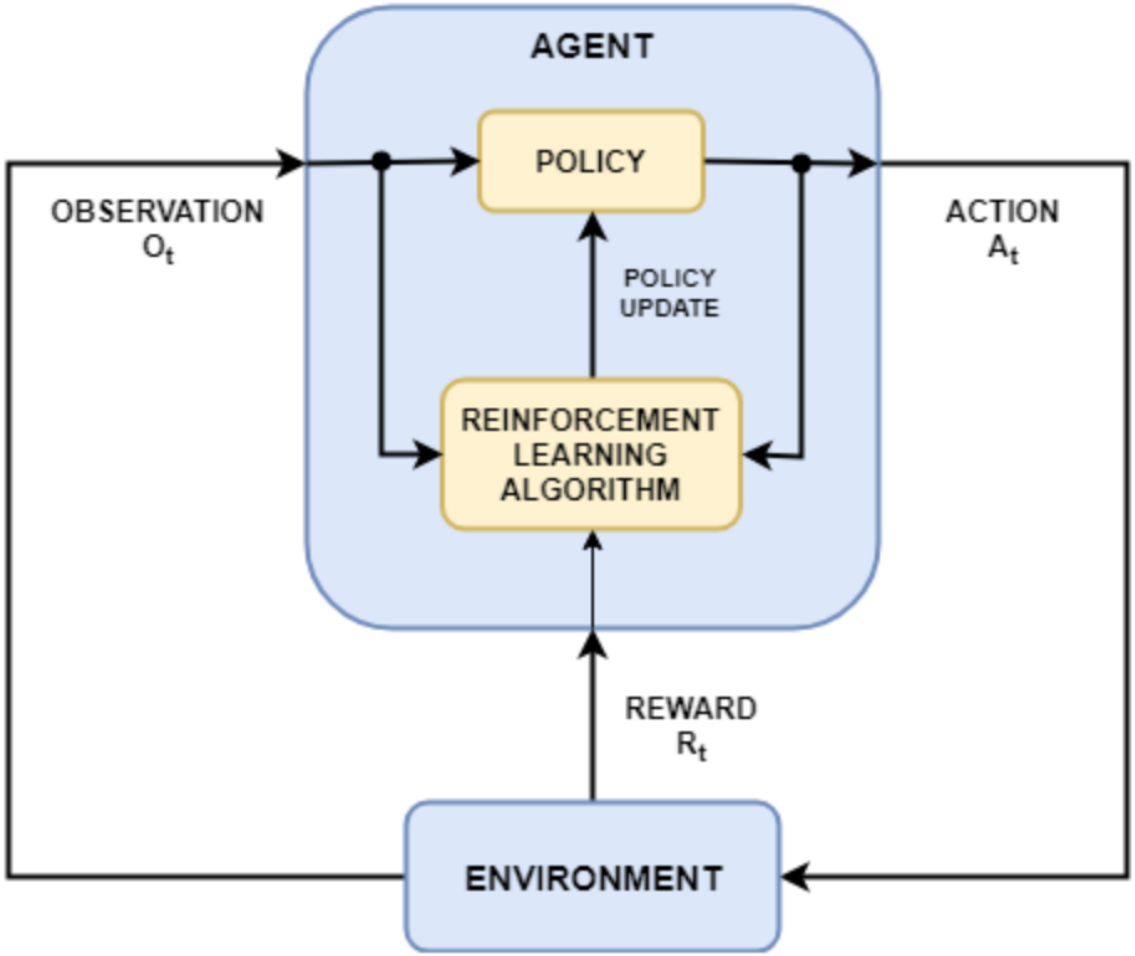
Reinforcement Learning

MATLAB Reinforcement Learning

- Reinforcement learning is a goal-directed computational approach where a computer learns to perform a task by interacting with an unknown dynamic environment.
- This learning approach enables a computer to make a series of decisions to maximize the cumulative reward for the task without human intervention and without being explicitly programmed to achieve the task.

web:https://www.mathworks.com/help/reinforcement-learning/ug/what-is-reinforcement-learning.html

MATLAB Reinforcement Learning



web:https://www.mathworks.com/help/reinforcement-learning/ug/what-is-reinforcement-learning.html

MATLAB Reinforcement Learning

- The goal of reinforcement learning is to train an agent to complete a task within an unknown environment.
- The agent receives observations and a reward from the environment and sends actions to the environment.
- The reward is a measure of how successful an action is with respect to completing the task goal.
- The agent contains two components: A policy and a learning algorithm.
- The policy is a mapping that selects actions based on the observations from the environment.
- .....

web:https://www.mathworks.com/help/reinforcement-learning/ug/what-is-reinforcement-learning.html

MATLAB Reinforcement Learning

- .....
- Typically, the policy is a function approximator with tunable parameters, such as a deep neural network.
- The learning algorithm continuously updates the policy parameters based on the action, observations, and reward.
- The goal of the learning algorithm is to find an optimal policy that maximizes the cumulative reward received during the task.
- Reinforcement learning refers to an agent learning the optimal behavior through repeated trial-and-error interactions with the environment without human involvement.

web:https://www.mathworks.com/help/reinforcement-learning/ug/what-is-reinforcement-learning.html

MATLAB Reinforcement Learning Workflow

The general workflow for training an agent using reinforcement learning includes the following steps:

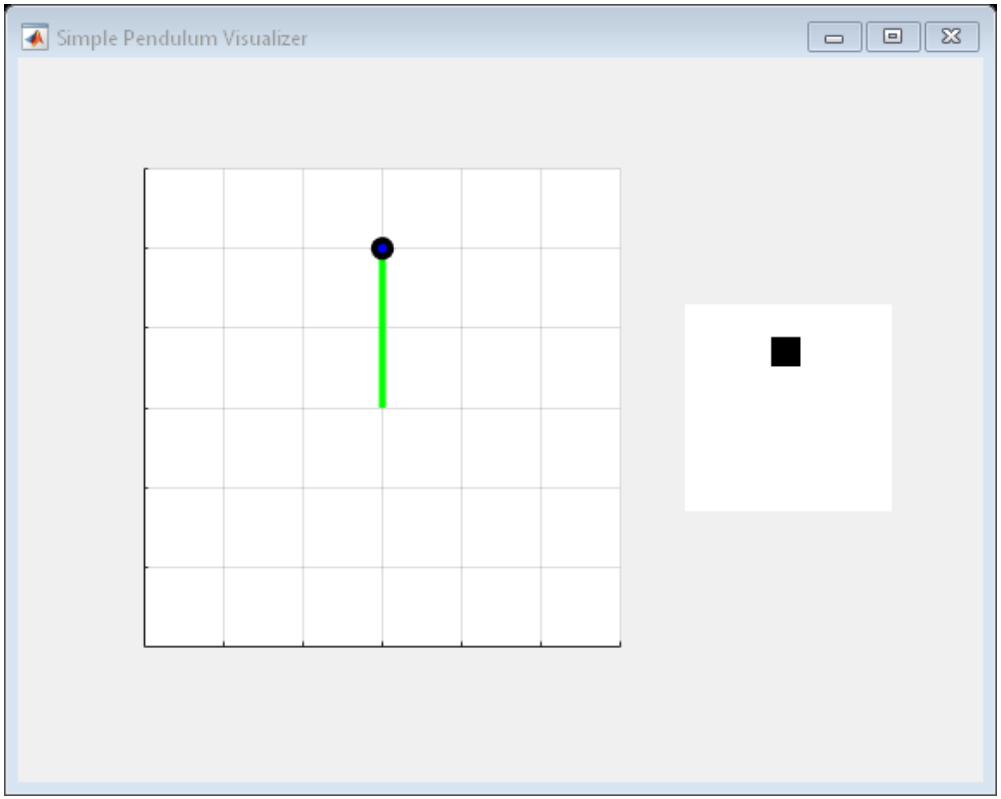
- Formulate Problem: Define the task for the agent to learn.
- Create Environment: Define the environment within which the agent operates.
- Define Reward: Specify the reward signal that the agent uses to measure its performance.
- Create Agent: Create the agent.
- Train Agent: Train the agent policy representation.
- Validate Agent: Evaluate the performance of the trained agent.
- Deploy Policy: Deploy the trained policy representation.

web:https://www.mathworks.com/help/reinforcement-learning/ug/what-is-reinforcement-learning.html

MATLAB Reinforcement Learning Workflow

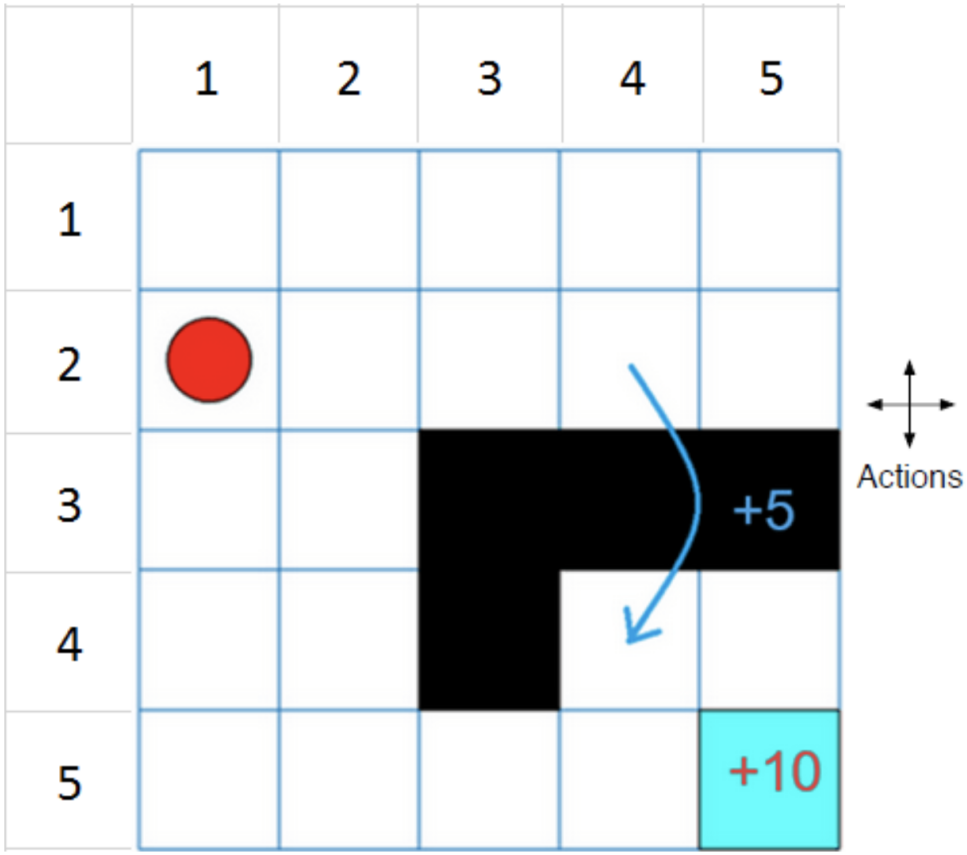
web:https://www.mathworks.com/help/reinforcement-learning/ug/what-is-reinforcement-learning.html

MATLAB Reinforcement Learning: Balance Pendulum



web:https://au.mathworks.com/help/deeplearning/ug/train-ddpg-agent-to-swing-up-and-balance-pendulum-with-image-observation.html

MATLAB Reinforcement Learning: Basic Grid World



web:https://www.mathworks.com/help/reinforcement-learning/ug/train-q-learning-agent-to-solve-basic-grid-world.html

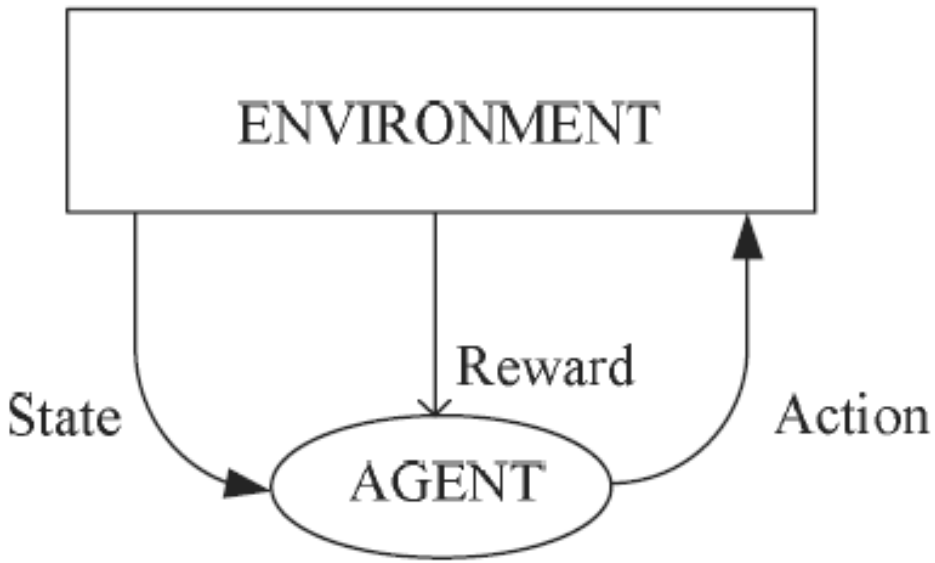
Policy and Value Iterations

- Reinforcement learning is learning what to do - how to map situations to actions — so as to *maximize a numerical reward.*
- Reinforcement learning receives *reward/penalty* or *trial/error* for its actions in trying to solve a problem.
- Reinforcement learning could *learn the best policy* and *maximize the total reward.*
- The *sequence of actions* has the maximum cumulative reward.
- For each policy π, there is a reward $v^{\pi}(s_t)$, we hope to find the *optimal policy*,

$$v^{*}(s_t) = \max_{\pi}(v^{\pi}(s_t)), \forall s_t$$

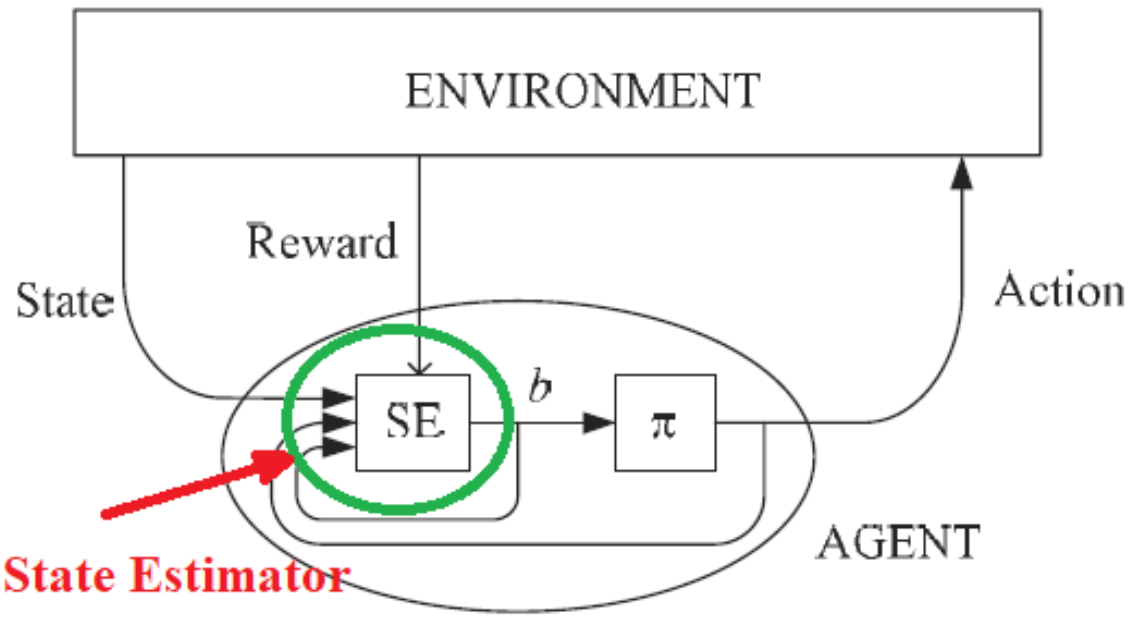
E. Alpaydn (2010) Introduction to Machine Learning (2nd Eition). The MIT Press.

Policy and Value Iterations



E. Alpaydn (2010) Introduction to Machine Learning (2nd Eition). The MIT Press.

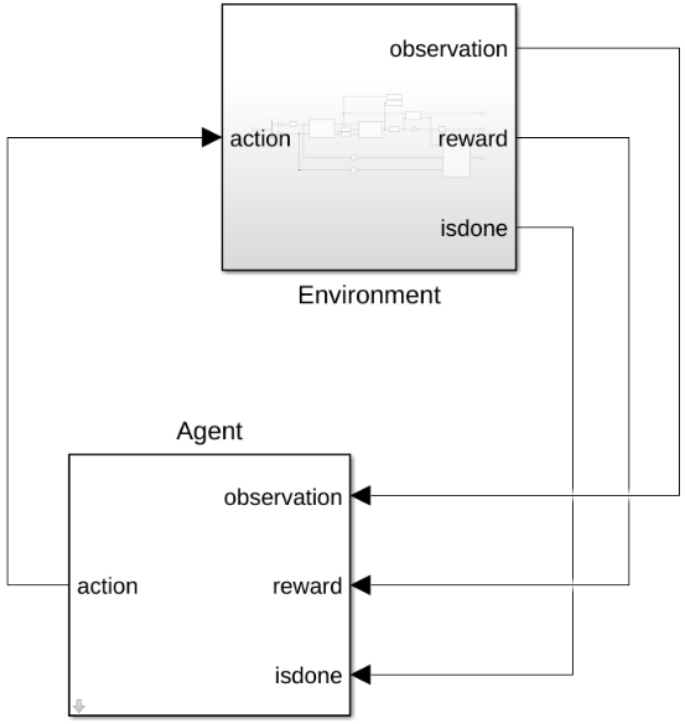
Policy and Value Iterations



Alpaydn, E. (2010) Introduction to Machine Learning (2nd Eition). The MIT Press.

MATLAB Reinforcement Learning

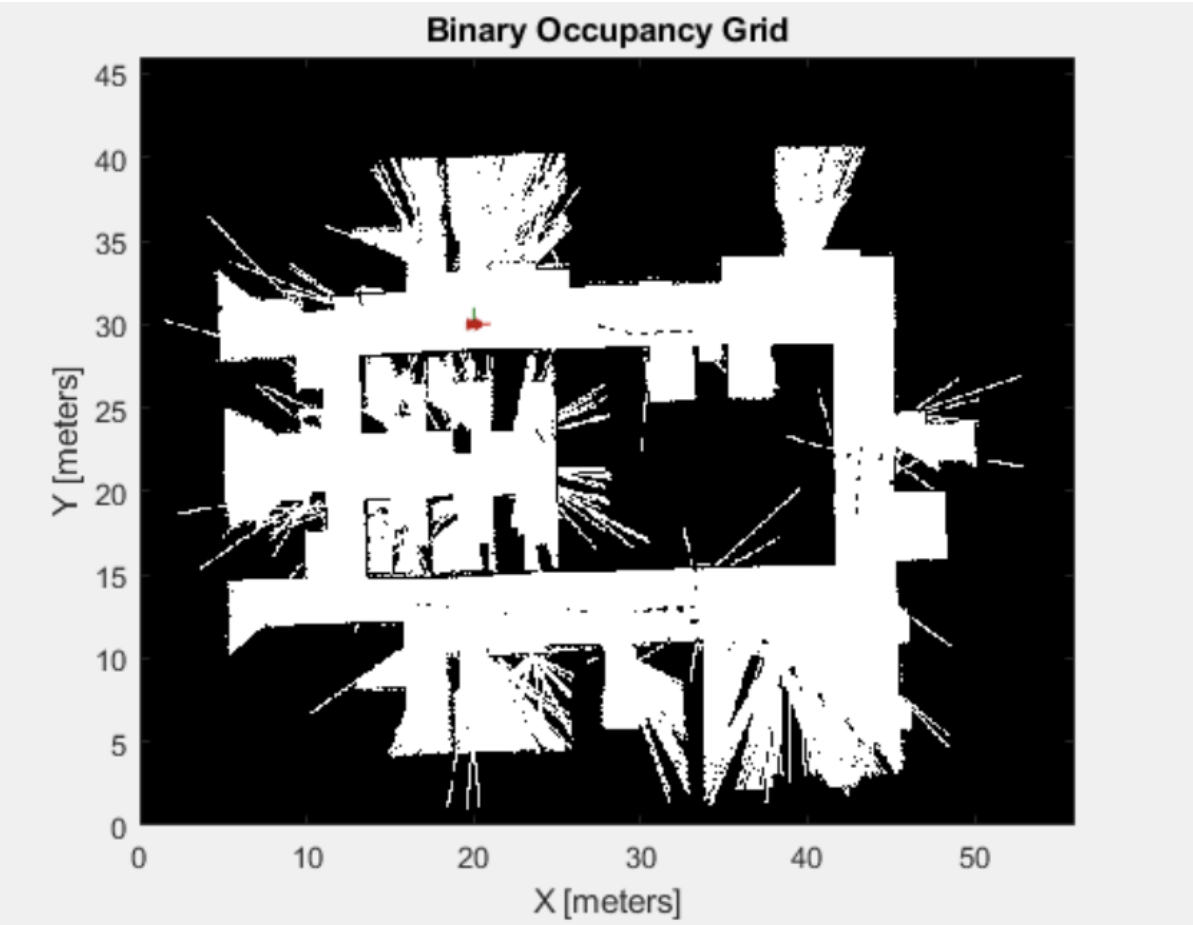
Avoid Obstacles Using Reinforcement Learning for Mobile Robots



https://au.mathworks.com/help/robotics/ug/avoid-obstacles-using-reinforcement-learning-for-mobile-robots.html

MATLAB Reinforcement Learning

Avoid Obstacles Using Reinforcement Learning for Mobile Robots



https://au.mathworks.com/help/robotics/ug/avoid-obstacles-using-reinforcement-learning-for-mobile-robots.html

Questions?

Reinforcement Learning

Questions?

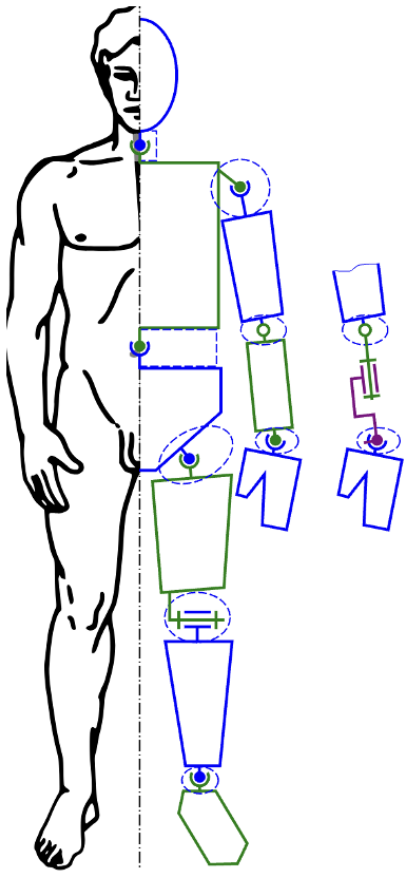
Reinforcement Learning enables a computer to make a series of decisions:

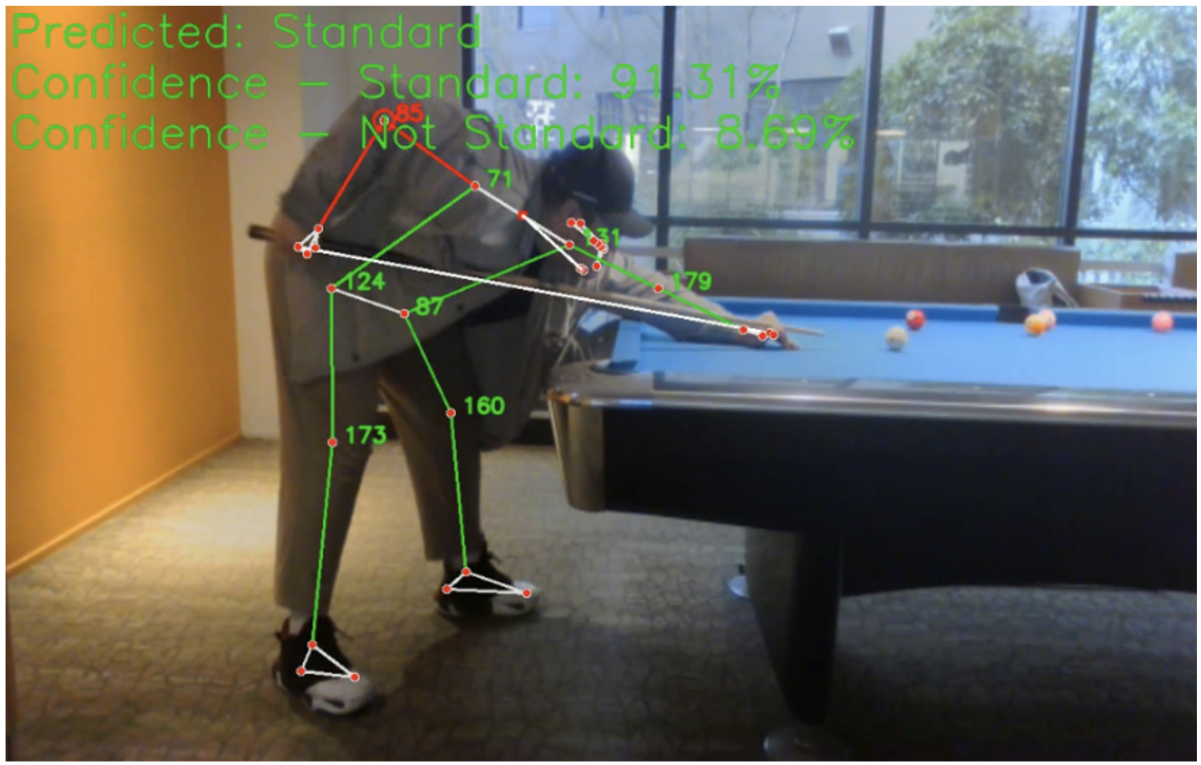
1. to maximize the cumulative reward for the task without human intervention and without being explicitly programmed to achieve the task.
2. to minimize the cumulative reward for the task without human intervention and without being explicitly programmed to achieve the task.
3. to maximize the cumulative reward for the task with human intervention and without being explicitly programmed to achieve the task.
4. to maximize the cumulative reward for the task with human intervention and being explicitly programmed to achieve the task.

which answer is right? ___

Questions?

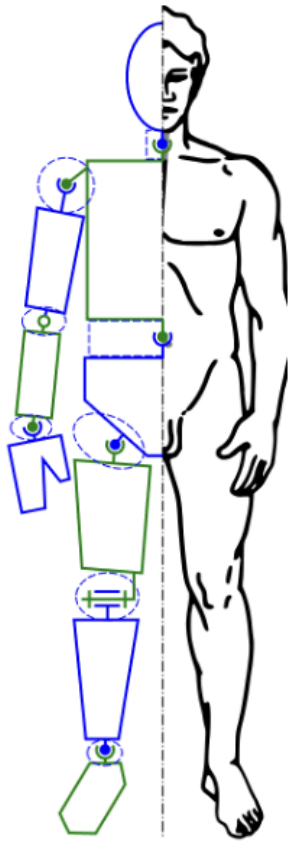
Imitation Learning

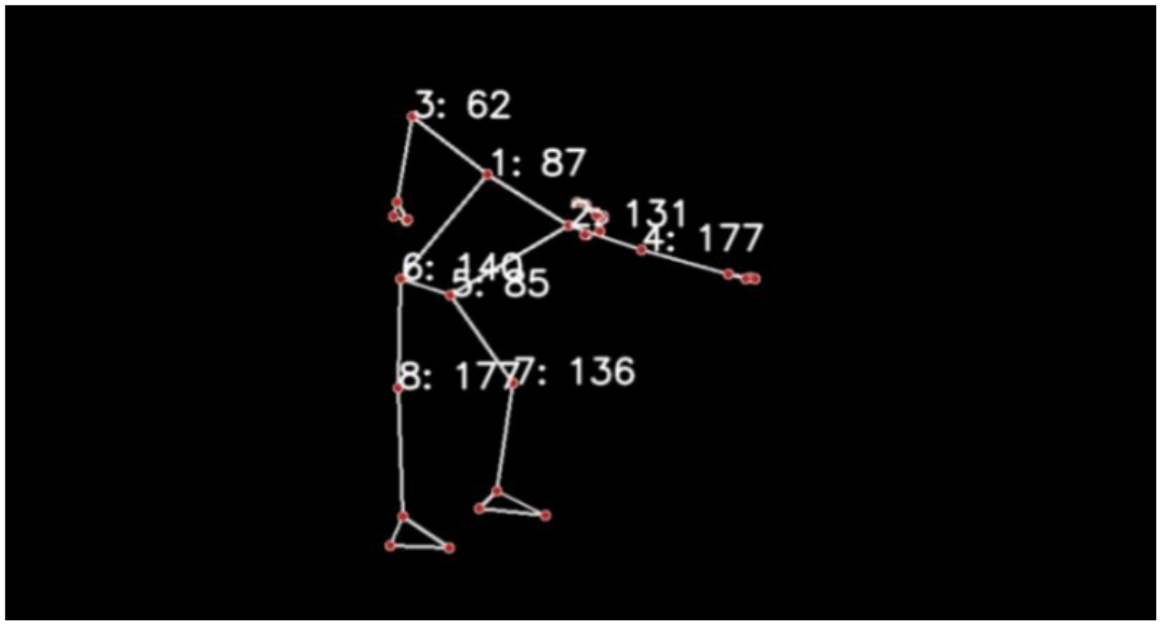




https://en.wikipedia.org/wiki/Imitation_learning

Imitation Learning





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

Imitation Learning

Imitation Learning

- In machine learning, an agent learns from trial and error within an environment, guided by a reward function.
- In imitation learning, the learning agent aims to mimic human behavior.
- In imitation learning, the agent learns from a dataset of demonstrations by an expert, typically a human.
- The goal is to replicate the expert's behavior in similar situations.

https://deepai.org/machine-learning-glossary-and-terms/imitation-learning

Imitation Learning

Imitation Learning

Imitation learning involves observing an expert performing a task and learning to imitate those actions. The three steps:

- *Data Collection*: An expert demonstrates the task to be learned. The actions and decisions of the expert are recorded as data.
- *Learning*: The collected data is then employed to train a machine learning model. The model learns a policy – a mapping from observations of the environment to actions.
- *Evaluation*: The trained model is tested in the environment to assess how well it performs the task compared to the expert. The goal is to minimize the difference between the expert's performance and the agent's performance.

https://deepai.org/machine-learning-glossary-and-terms/imitation-learning

Imitation Learning

Imitation Learning

There are two main approaches to imitation learning:

- *Behavioral Cloning*: The model is trained in a supervised learning fashion using the state-action pairs from the expert's demonstrations.
- *Inverse Reinforcement Learning (IRL)*: It aims to learn the underlying reward function that the expert seems to be maximizing. This approach can generalize better to unseen states.

https://deepai.org/machine-learning-glossary-and-terms/imitation-learning

Imitation Learning

Challenges

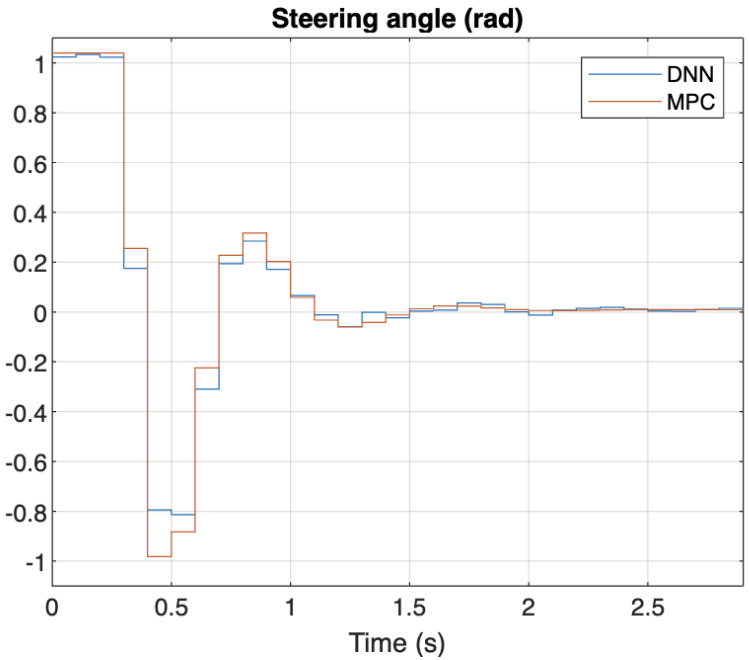
- *Data Quality*: The quality of the learned policy is highly dependent on the quality of the demonstrations.
- *Distribution Shift Problem*: The agent may encounter states that were not covered in the training demonstrations, leading to uncertain behavior.
- *Scalability*: Collecting expert demonstrations can be expensive and time-consuming, especially for complex tasks.
- *Generalization*: The ability of the agent to generalize the learned behavior to new situations is a critical challenge, especially in dynamic and unpredictable environments.

https://deepai.org/machine-learning-glossary-and-terms/imitation-learning

Imitation Learning

MATLAB: Imitation learning for Lane Keeping

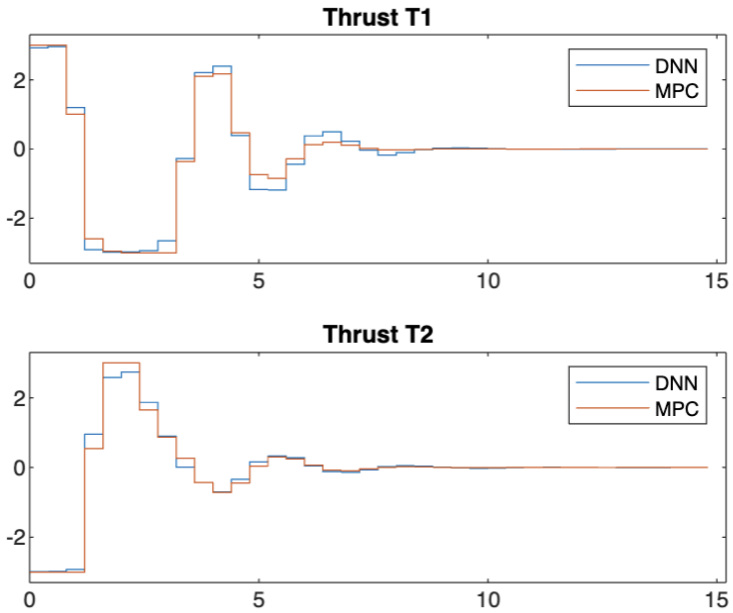
The deep neural network successfully imitates the behavior of the Model Predictive Controller (MPC). The vehicle state and control action trajectories for the controller and the deep neural network closely align.



https://au.mathworks.com/help/reinforcement-learning/ug/imitate-mpc-controller-for-lane-keeping-assist.html

MATLAB: Imitate for Flying Robot

The deep neural network successfully imitates the behavior of the NLMPC controller. The flying robot states and control action trajectories for the controller and the deep neural network closely align.



https://au.mathworks.com/help/reinforcement-learning/ug/imitate-nonlinear-mpc-controller-for-flying-robot.html

Questions?

Imitation Learning

Questions?

Imitation learning includes:

1. Inverse reinforcement learning
2. Reinforcement learning
3. Reward learning
4. Feedback learning

which answer is right?___

Questions?

Learning Objectives

- Derive solutions for particular robotic vision and visual control tasks characterised by specifics of image data and deep learning algorithms.
- Critically evaluate the performance of robotic vision with deep learning algorithms, bench mark data, performance measures, and ways to define ground truth.
- Examine opportunities of using robotic vision as a part of complex robotic systems and applications.