Take me home

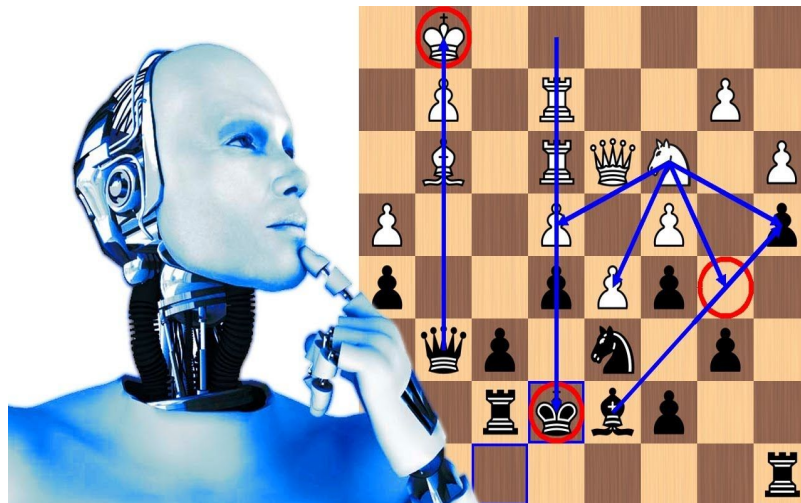## Land the spacecraft at target pad

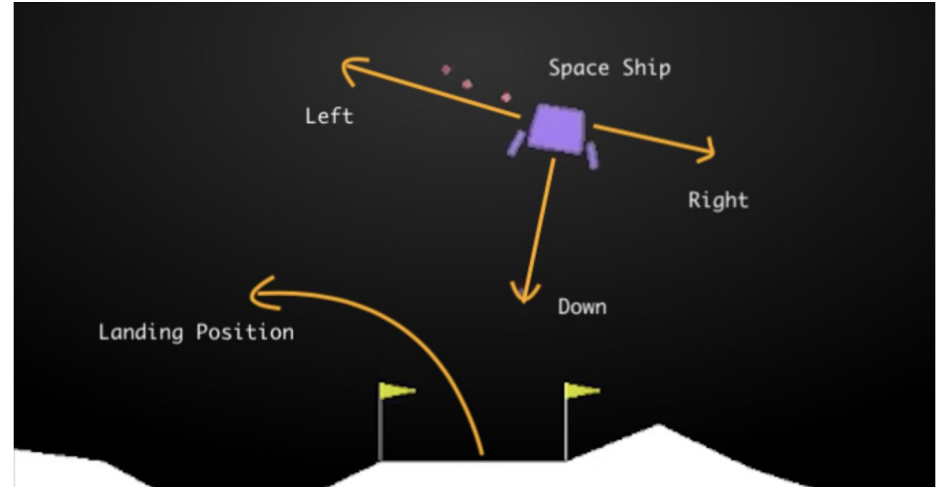Sunil Kumar J S

Human vs Computer

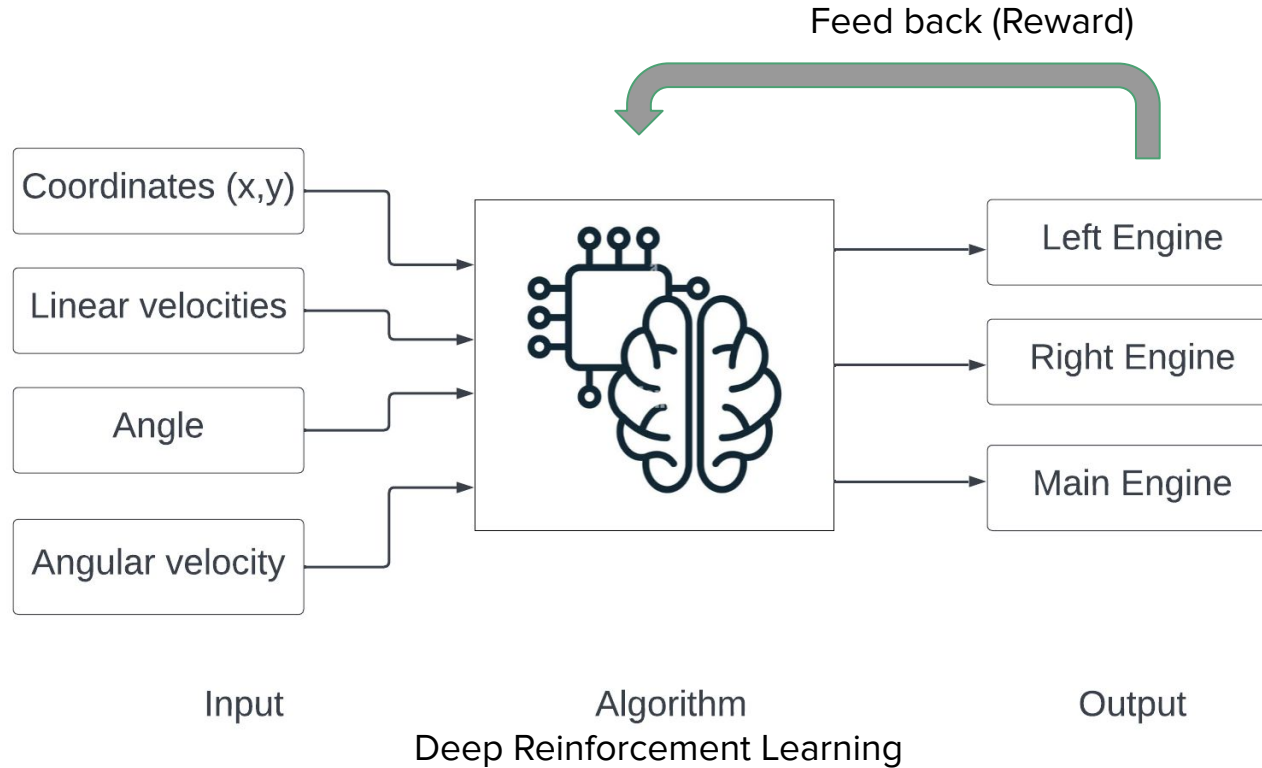Magnus Carlsen

2864

AlphaZero

4680

Problem Statement

Spacecraft starts at top center with random initial force applied. It needs to be landed between the flags (home) using three engines.

Data : open ai gym

Solution

Feed back (Reward)

Input	Algorithm	Output
Coordinates (x,y)		Left Engine
Linear velocities		Right Engine
Angle		Main Engine
Angular velocity		

Deep Reinforcement Learning

Deep Q learning Algorithm

Algorithm 1 Deep Q-learning with Experience Replay

Initialize replay memory $\mathcal{D}$ to capacity N
Initialize action-value function Q with random weights
for episode $= 1, M$ **do**
    Initialise sequence $s_1 = \{x_1\}$ and preprocessed sequenced $\phi_1 = \phi(s_1)$
    for $t = 1, T$ **do**
        With probability ϵ select a random action a_t
        otherwise select $a_t = \max_a Q^*(\phi(s_t), a; \theta)$
        Execute action a_t in emulator and observe reward r_t and image x_{t+1}
        Set $s_{t+1} = s_t, a_t, x_{t+1}$ and preprocess $\phi_{t+1} = \phi(s_{t+1})$
        Store transition $(\phi_t, a_t, r_t, \phi_{t+1})$ in $\mathcal{D}$
        Sample random minibatch of transitions $(\phi_j, a_j, r_j, \phi_{j+1})$ from $\mathcal{D}$
        Set $y_j = \begin{cases} r_j & \text{for terminal } \phi_{j+1} \\ r_j + \gamma \max_{a'} Q(\phi_{j+1}, a'; \theta) & \text{for non-terminal } \phi_{j+1} \end{cases}$
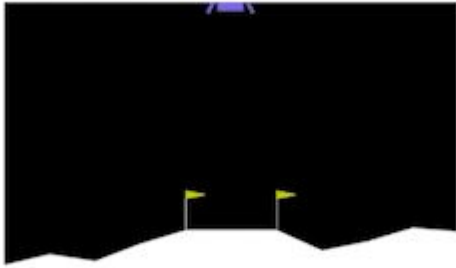   Perform a gradient descent step on $(y_j - Q(\phi_j, a_j; \theta))^2$ according to equation 3
    end for
end for

Training Progression

Naive Agent

Moderately trained Agent

Fully Trained Agent

Comparison of Training Progression :