Ball Kicking for Reinforcement Learning on Poppy Humanoid

Initial Presentation of Applied Reinforcement Learning

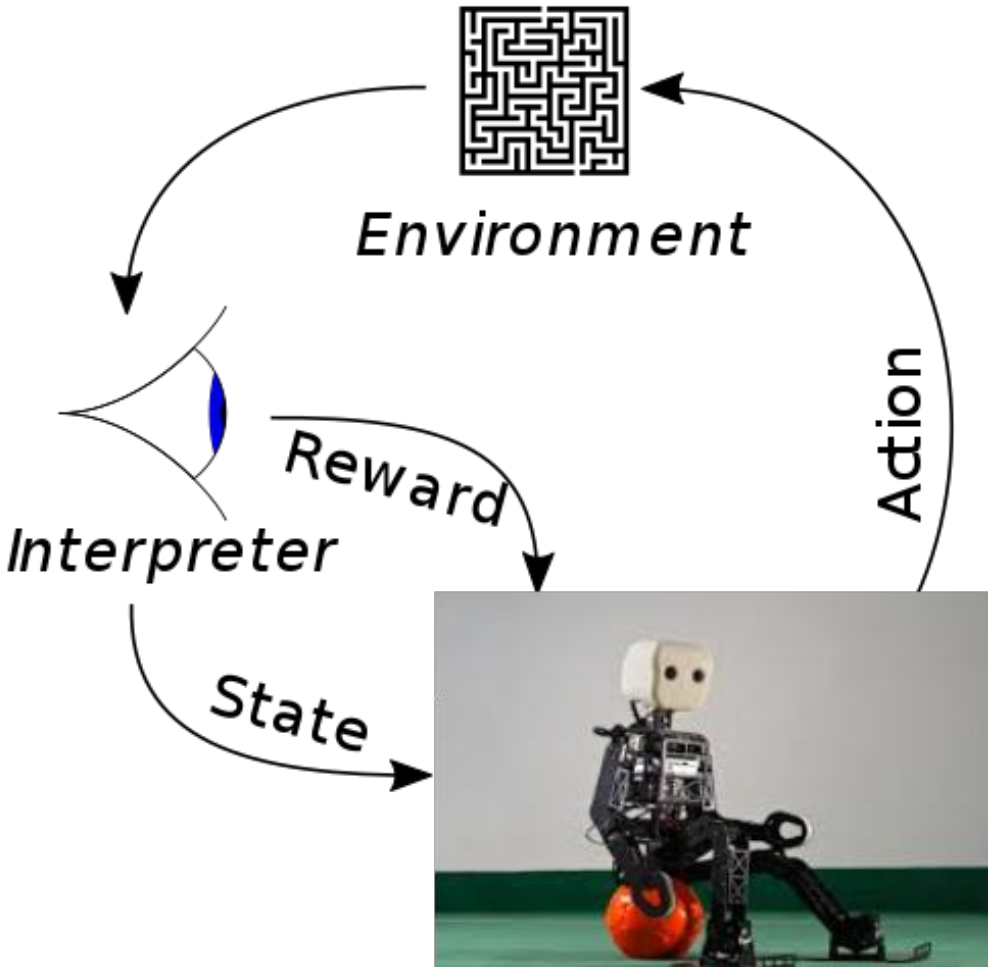
Yinan Shi, Fengze Han, Hao Xu, Jiaojiao Ye

Institute for Data Processing, Technical University of Munich

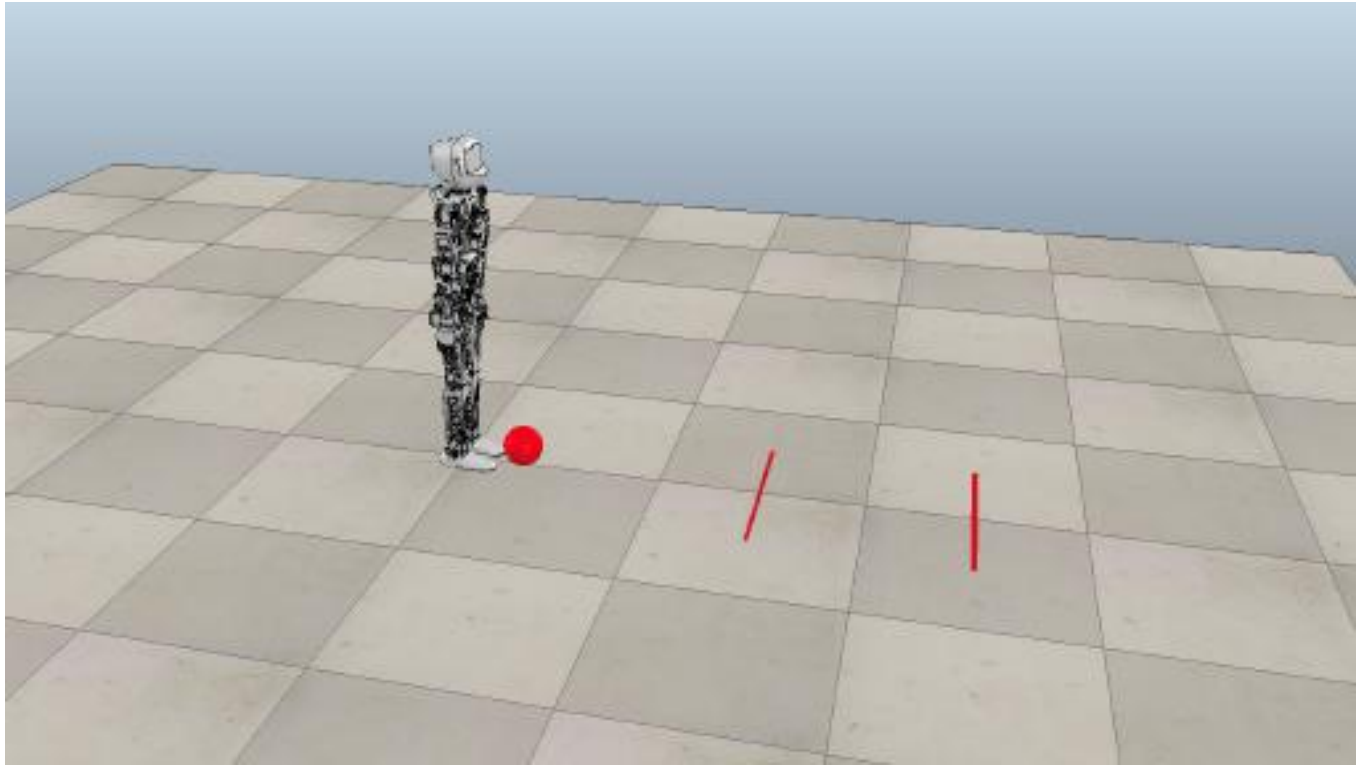
May 2, 2018

Table of contents

Motivation

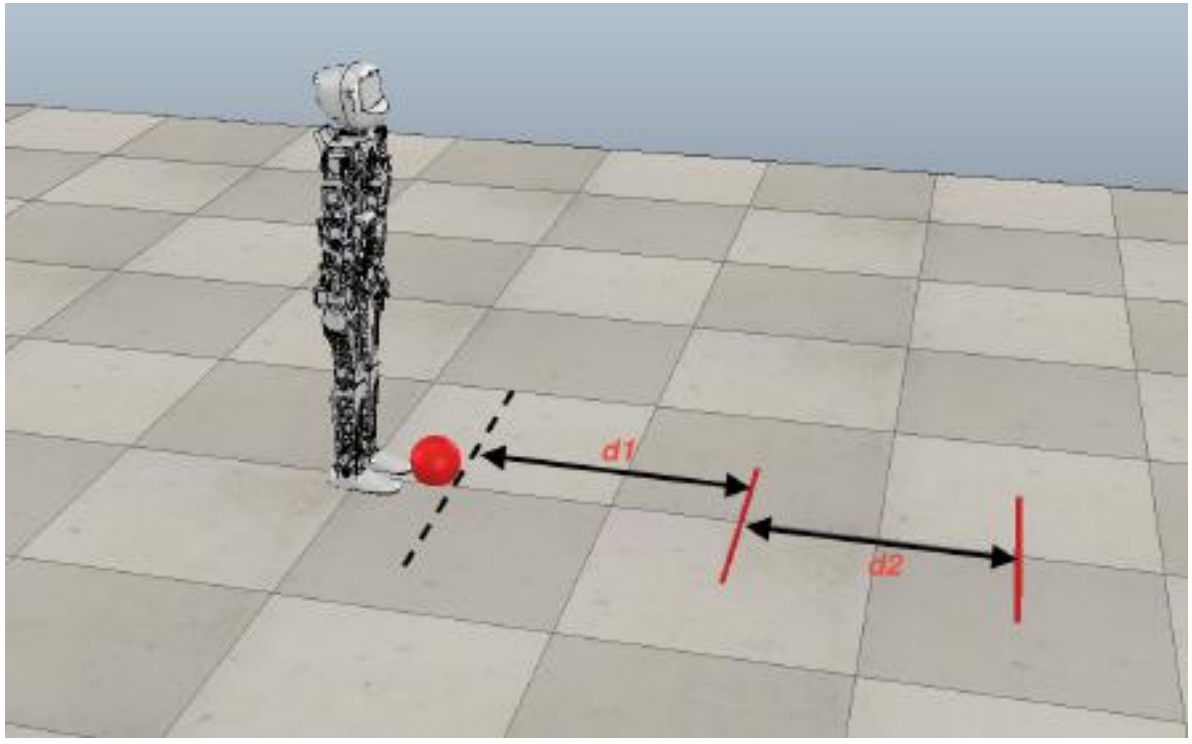


Problem Description



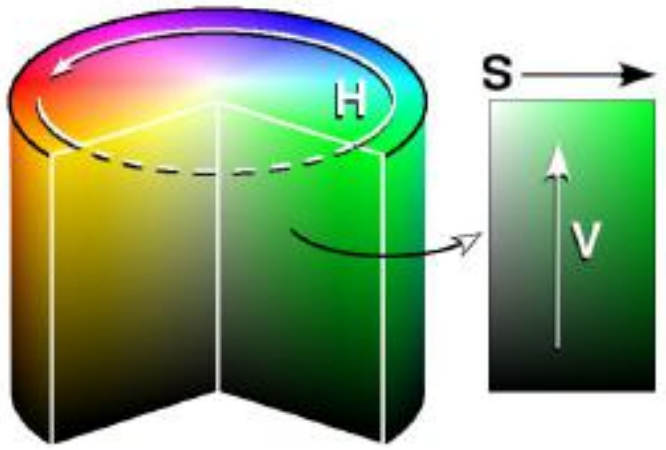
The goal is to train a poppy humanoid that sits on a chair to kick the ball into the target region.

States



- Ball and lines detection using real-time images
- States are constructed by Pixel distance d_1 and d_2 in form(d_1,d_2)
- Goal state is kicking the ball into the region between two red lines

Image processing and Object detection



- H – Hue(Dominant Wavelength)
- S – Saturation (Purity / shades of the color)
- V – Value(Intensity)

After get the pixels of red in the HSV color space, use Hough Transform to detect lines and circle(ball).

Action

- Motor angle in knee starting from an initial value θ_0
- Angular velocity control, e.g. a 2-dimensional vector of goal position and time duration: $[\theta_t, t]^T$
- Constraints on values and sensitivity need to be studied
- After test on real robot, action space and how to do discretization will be determined

Reward

- Most critical part of a reinforcement learning scenario
- Assign ’-10’ when the episode terminates outside the desired region; assign ’+10’ otherwise
- Both values will be tuned in practice to find out the best choice

Algorithm

TD Learning in Q function

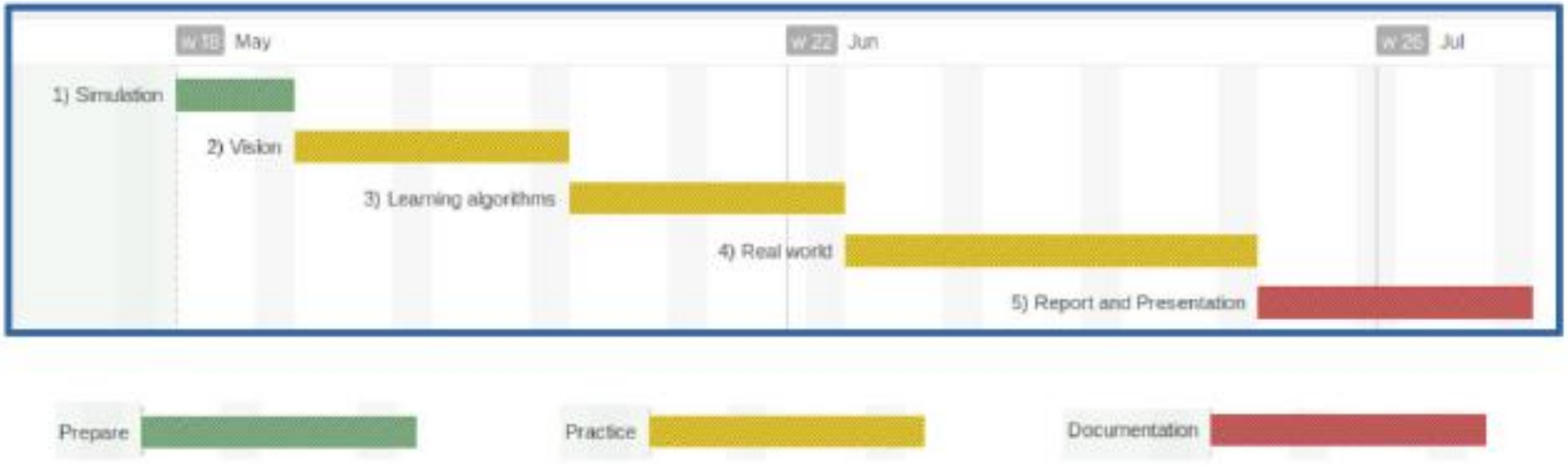
SARSA

$$Q(s,a) \leftarrow Q(s,a) + \alpha(r(s,a,s') + \gamma Q(s', \pi(s')) - Q(s,a))$$

Q-Learning

$$Q(s,a) \leftarrow Q(s,a) + \alpha(r(s,a,s') + \gamma \min_{a' \in A(s')} Q(s',a') - Q(s,a))$$

Time Plan


w 18
May
w 22
Jun
w 26
Jul
1) Simulation
2) Vision
3) Learning algorithms
4) Real world
5) Report and Presentation
Prepare
Practice
Documentation