Intro to RL and MDPs - Personal Notes

Diego Andrés Gómez Polo

February 2024

Contents

1 Introduction

For these note i will be guiding myself with [Montague, 1999]

1.1 RL vs Supervised Learning vs Unsupervised Learning

Hedonistic learner (an actor that wants something from an environment interactionz) These two characteristics—trial-and-error search and delayed reward—are the two most important distinguishing features of reinforcement learning.

Markov decision processes are intended to include just these three aspects—sensation, action, and goal—in their simplest possible forms without trivializing any of them. Any method that is well suited to solving such problems we consider to be a reinforcement learning method.

In supervised learning given a state (data) you have knowledge about the correct action the system must take, usually to give a label to the data, and try to extrapolate this correct actions to other uknown states. In RL you do not know the correct action given a state, rather you are given a certain numerical reward and are asked to maximized the long term reword.

In a biological system, we might think of rewards as analogous to the experiences of pleasure or pain.

Exploration vs exploitation (mejor vale malo conocido que bueno por conocer? no siempre...)

Hierarchical planning? Hierarchical exploration?

How you think of an already learned hand movement? Articulation by articulation or as a whole? How do yo learn a new hand or body movement, very specific, practice its somposing body steps until it becomes a "larger" mental image

Touch is incredible dense, a lot of information about force vectors on many specific locations.

Intelligence emerges in higher level plans

1.2   Markov Decision Processes (MDPs)

References

[Montague, 1999] Montague, P. R. (1999). Reinforcement learning: an introduction, by sutton, rs and barto, ag. *Trends in cognitive sciences*, 3(9):360.