note set -1

Introduction to Reinforcement Learning ISAT 670

"Patterns & Learning"

- In this class we will study an important topic "Reinforcement Learning"
- It is a branch of AI / machine learning that has been a key part of some of the interesting advances in A/I and deep learning in recent years

Here is a paraphrased definition – (synopsis)

Reinforcement Learning is the science of decision making. An agent tries to learn the optimal behavior in an environment in order to obtain the maximum reward. To do this the agent interacts with the environment, and observes the results of its interactions

This is similar to children exploring the world around them and learning the actions that help them acheive their goal.

Since there is no supervisor the learner must discover the sequence of actions that maximize the reward (acheivement of the goal state). The process of discoveng is a mix of exploration (trial and error) and exploitation (making decisisions based on the current state of knowledge)

How does it work?

- The agent explores an unknown environment in an effort to acheive a goal state.
- Cumulitive reward is central to Reinforcement Learning
- The formal framework of RL is based on Markov Decision Processes (MDP)

Elements of a RL system are:

- the agent
- the environment
- the rules / policy the agent follows to take actions
- the reward observed for taking various actions

RL algorithms can be

Model Free - various actions of the agent in the environment are used to find an optimal policy

or

Model Based - optimal policy is derived without necessarily exploring states of an environment

Examples of Reinforcement Learning

Robotics - in unknown environments, pre programmed behavior sometimes (lots of times) falls short. Think of an environment with obstacles whose location is not known. RL provides a way to find paths between locations for the robot to follow.

Alpha Go - RL helped a computer (2016) beat the best go player. It learned by playing thousands of games with experienced players

go is a 3000 year old Chinese board game with 10^{270} possible board combinations

The latest version can learn by playing against itself

Autonomous Driving - RL can help with path planning and motion prediction

Benefits of RL

- focuses on problem as a whole - RL does not divide a problem into subtasks, it works to maximize long term reward
- Does not need a seperate data collection step - learns by direct interaction with environment
- works in dynamic uncertain environments

Challenges

- RL agent needs lots of experience
- Delayed Rewards - Jare used for discovery of the optimal policy, they can sometimes make it hard to find the optimal policy
- Sometimes it is hard to understand why the agent does what it does

RL vs Supervised Learning

Supervised learning requires a labelled data set. Once this data is established it is fed to the learning algorithm. The learning algorithm's goal is to generalize, interpelate, etc, so it can make predictions for data not included in the training set.

RL does not require labelled data, it acquires this info by interacting with the environment.

Is RL unsupervised learning? No, unsupervised learning's objective is to determine the structure of the data, (Think clustering)