



Grade 11/12 Math Circles

March 22, 2023

Reinforcement Learning - Solutions

Problem Set Solutions

1. Tom Mitchell provides a modern definition of Machine learning: “A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P , if its performance at tasks in T , as measured by P , improves with experience E .” Use this definition to classify T , P and E for playing chess.

Solution: T is the task of playing chess

P is the probability that the program will win the next game

E is the experience of playing many games of chess

2. Consider the following experiment in evolution. We are interested in the evolution of a gene by the name mTOR. The gene occurs in two variants T and S and plays an important role in determining the height of an individual. Each individual has a pair of this gene, either TT (tall) or TS (medium height) or SS (short). In the case where the individual is of medium height, the order of gene pairing is irrelevant (TS or ST is same). In evolution, an offspring inherits a pair of gene from each of his/her biological parents with equal probability. Thus if one partner is tall(TT) and the other partner is medium(TS), the offspring has $1/2$ probability of being tall or $1/2$ probability of being medium. We shall start with an individual of an arbitrary but fixed trait(TT or TS or SS). The other partner is of medium height (ST or TS). The offspring of such partners again finds a medium height partner. This pattern repeats for a number of generations, wherein the resultant offspring always has a medium height partner. Write out the states and transition probabilities of this Markov process.

Solution:

$$\text{States} = \{TT, SS, ST\}, \mathcal{P} = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{bmatrix}$$

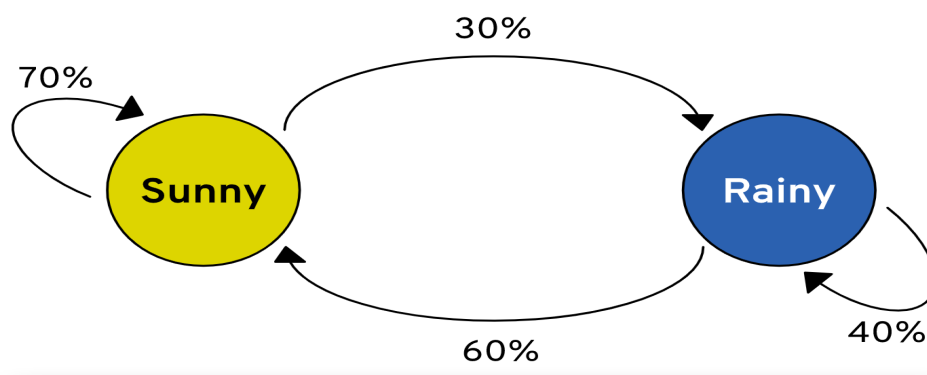


Figure 1: Simple two state Markov chain

3. Write out the states and transition probability matrix corresponding to Fig. 1

Solution: States = {Sunny, Rainy},

$$\mathcal{P} = \begin{bmatrix} 0.70 & 0.30 \\ 0.60 & 0.40 \end{bmatrix}$$