

COMP-424 - Assignment 3

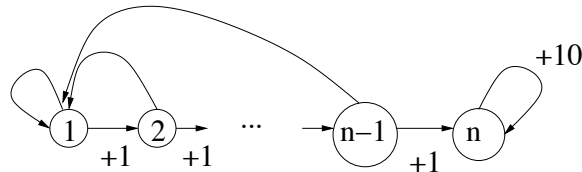
Posted Monday April 8, 2013

Due Monday April 15, 2013

No penalties until Monday April 22, 2013

1. [15 points] Markov Decision Processes

Consider the n -state MDP in the figure below. In state n there is just one action that loops around forever, and collects a reward of $+10$. In all the other states there are two actions: float, which moves deterministically one step to the right, and reset, which deterministically goes back to state 1. There is a reward of $+1$ for a float and 0 for reset. The discount factor is $\gamma = \frac{1}{2}$.



(a) [5 points] What is the optimal policy?

(b) [10 points] Compute the optimal value function, $V^*(k)$ and the optimal action, for all $k = 1, \dots, n$.

2. [10 points] Gradient descent

Consider a different kind of computing unit for a neural network. Given an input vector $\mathbf{x} = \langle 1, x_1, \dots, x_n \rangle$ and a weight vector \mathbf{w} , the output of the unit will be:

$$o = e^{-(\mathbf{w} \cdot \mathbf{x})}$$

Derive a gradient descent training rule for a single unit of this type

3. [15 points] Perceptrons.

(a) [5 points] Draw a perceptron that computes the Boolean implication function, $a \Rightarrow b$ (specifying the inputs and the weights)

(b) [10 points] Suppose you want to represent the function $a \Rightarrow (b \wedge c)$. Show either a perceptron or a network of perceptrons as small as possible to compute this function.

4. [20 points] Bayes nets

The starship Enterprise is being attacked by the Romulans, and Data the android is trying to save it. He knows that the Romulans have two types of torpedoes. Type A will destroy the Enterprise with 80% probability in one shot, if the Enterprise is near the Romulan ship, and with 30% probability if it is far. Type B will destroy The Enterprise with 60% probability if it is near, and with 10% probability if it is far. The Romulan ship can carry at most 1 missile of each type. The Enterprise

can only destroy the Romulans if it is near their ship. Data has an imprecise sensor for detecting if missiles are on board the Romulan ship, and the sensor is telling him that there is 1 type A missile and no type B missile. However, the sensor has only a 70% chance of success. If it fails, it returns a uniformly random reading for each of the missiles.

- (a) [10 points] Represent this problem as a Bayesian network
- (b) [10 points] What are the probabilities of the Enterprise being destroyed if it gets near the Romulan ship, conditioned on the observed sensor readings?

5. [20 points] **Maximum likelihood learning**

The normal distribution is one of the most important probability distributions, used widely in many scientific disciplines.

- (a) [10 points] The univariate normal distribution has the following probability density function:

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

where μ and σ are the mean and variance parameters. Suppose that you do not know μ and σ , but you have examples x_1, \dots, x_n drawn i.i.d. from the distribution in which you are interested. Write down the log likelihood of the data, take the derivatives with respect to μ and σ , set them to 0 and derive the expression of the maximum likelihood estimators for these parameters, denoted μ^n and σ_n .

- (b) [10 points] Suppose that you now get a new data point x_{n+1} . Write down an update rule which computes μ_{n+1} and σ_{n+1} from μ_n, σ_n and x_{n+1} (without involving any of the previous data points x_1, \dots, x_n)

6. [10 points] **Neural networks**

In class we showed how to train the weights of a neural network using gradient descent. Could we use simulated annealing for the same purpose? If your answer is yes, explain how this would be implemented, and whether this would work better or worse than plain gradient descent. If your answer is no, explain why not.

7. [10 points] **Problem formulation**

You have been hired by a large online retail company who wants to send advertisements to customers on social media. They have the purchasing history of every user (date, prices, items, shipping address) for the last 5 years. They do not want to overwhelm users with the ads. Describe precisely *one* AI method that you could apply to this problem, and how this would be done. Eg. for a search problem, states, actions... ; for a learning problem, data encoding, learning method...