NAME: ...KEY ....
CSCI E-89c Deep Reinforcement Learning
Part I of Assignment 10

Suppose each state $s \in \mathcal{S}$ of the Markov Decision Process can be represented by a vector of 2 real-valued features: $\mathbf{x}(s) = (x_1(s), x_2(s))^T$.

Given some policy π, suppose we model the state value function $v_\pi(s)$ with a *fully connected feedforward neural network* (please see the table below) which has two inputs ($x_1(s)$ and $x_2(s)$), one hidden layer that consists of two neurons (u_1 and u_2) with ReLU activation functions, and one output ($\hat{v}(s, \mathbf{w})$) with the ReLU activation function.

The explicit representation of this network is

input layer	hidden layer	output layer
x_1	$u_1 = f(w_{01}^{(1)} + w_{11}^{(1)} x_1 + w_{21}^{(1)} x_2)$	$\hat{v} = f(w_0^{(2)} + w_1^{(2)} u_1 + w_2^{(2)} u_2)$
x_2	$u_2 = f(w_{02}^{(1)} + w_{12}^{(1)} x_1 + w_{22}^{(1)} x_2)$	

Here, $f(x)$ denotes the rectified linear unit (ReLU) defined as follows:

$$f(x) = \begin{cases} x, & \text{if } x \geq 0, \\ 0, & \text{if } x < 0. \end{cases}$$

(a) Sketch the graphical representation of this neural network. Indicate values $\hat{v}$, x_1, x_2 and u_1, u_2 next to the corresponding neurons in the graph. Also, indicate weights next to corresponding connections.

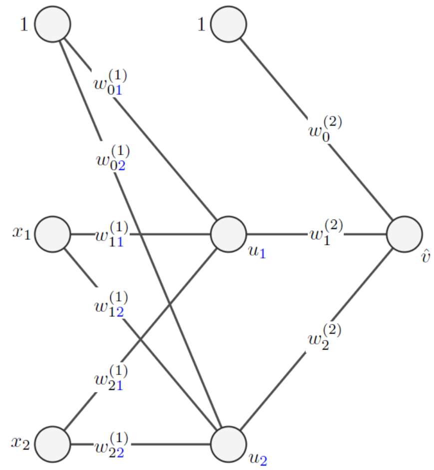
(b) Assume we trained the network (by minimizing the loss function) and came up with the following weights $\mathbf{w}$:

hidden layer	output layer
$w_{01}^{(1)} = -1.2, w_{11}^{(1)} = 0.1, w_{21}^{(1)} = 0.5$	$w_0^{(2)} = 0.2, w_1^{(2)} = -0.8, w_2^{(2)} = 1.2$
$w_{02}^{(1)} = 0.9, w_{12}^{(1)} = 0.8, w_{22}^{(1)} = -0.3$	

If the features of some state s are $x_1(s) = 1.3$ and $x_2(s) = 0.7$, what is the approximation $\hat{v}_\pi(s, \mathbf{w})$ of the state value $v_\pi(s)$ obtained with this neural network?

SOLUTION:

(a) Graphical representation of this neural network:



(b) We have

$$u_1 = f(w_{01}^{(1)} + w_{11}^{(1)} x_1 + w_{21}^{(1)} x_2) = f(-1.2 + 0.1 \cdot 1.3 + 0.5 \cdot 0.7) = f(-0.72) = 0$$
$$u_2 = f(w_{02}^{(1)} + w_{12}^{(1)} x_1 + w_{22}^{(1)} x_2) = f(0.9 + 0.8 \cdot 1.3 - 0.3 \cdot 0.7) = f(1.73) = 1.73$$

then

$$\hat{v} = f(w_0^{(2)} + w_1^{(2)} u_1 + w_2^{(2)} u_2) = f(0.2 - 0.8 \cdot 0 + 1.2 \cdot 1.73) = f(2.276) = 2.276.$$