

14. The Bellman Equation :

Theme:

Date: / /

Sat Sun Mon Tue wed Thu Fri




S = State

R - Reward

a = Action

γ / gamma = Discount factor .




When Agent is here and goes right it gets +1, so it checks how it earn to this position. and gives it a value of 1 and keeps doing this, until it reaches starting position.

$V=1$	$V=1$	$V=1$	
$V=1$			
$V=1$			
			

Agent starts from here

$R=+1$

$R=-1$

	$V=1$	$V=1$	
$V=1$			
$V=1$			

But what about then when the agent is in this position ? DOES NOT WORK !

That's when the Bellman equation comes in !

$$V(s) = \max_a (R(s,a) + \gamma V(s'))$$

*

V = Value of being in a certain state .

s = current state on any given state .

s' = the following state . The state you will end up in after this state .

R = reward

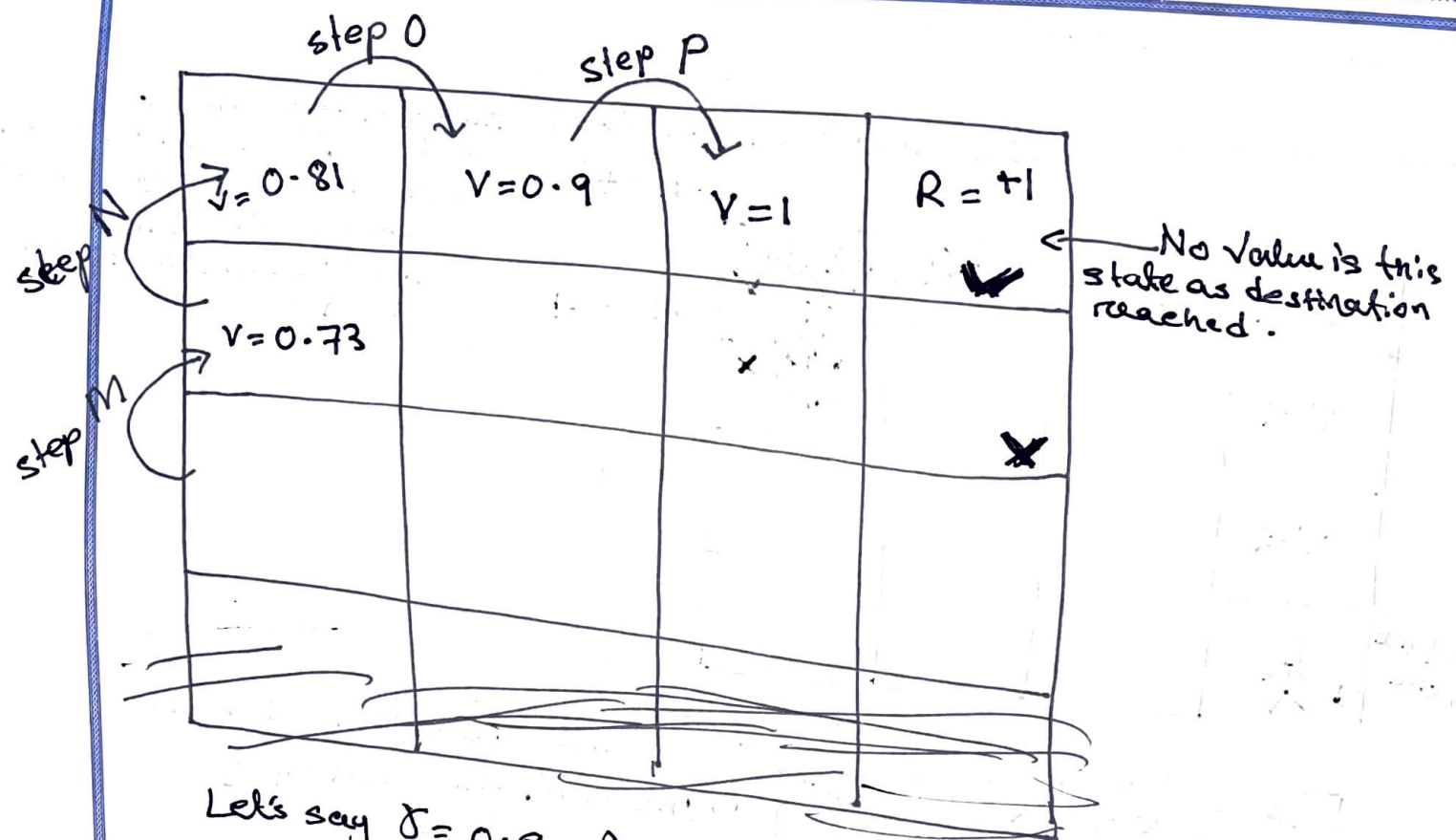
\max_a = maximum value based on an action .

$$V(s) = \max_a (R(s,a) + \gamma V(s'))$$

Theme:

Date:

☐ Sat ☐ Sun ☐ Mon



Let's say $\gamma = 0.9$, $R = 0$, $V = 1$

For step P : $V(s) = \max_a (0 + 0.9(\text{value of next stage}))$
 $\Rightarrow V(s) = \max_a (0 + 0.9(1))$
 $\Rightarrow V(s) = 0.9$

For step 0 :

$$V(s) = \max_a (0 + 0.9(0.9))$$

$$V(s) = 0.81$$

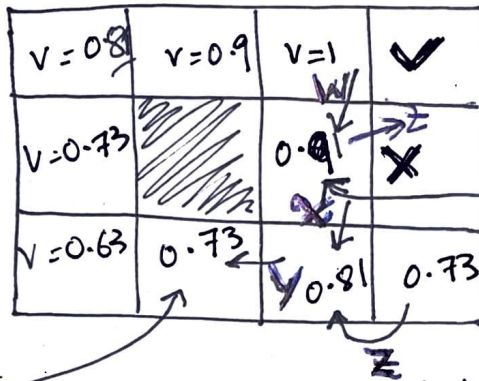
For step N : $V(s) = (0 + 0.9(0.81))$
 $= 0.73$

For M : $V(s) = (0 + 0.9(0.73))$
 $= 0.66$

Theme:

(Q) How does the Discount factor work?

(Ans) It discounts the value (V) of the state (s) as you are further away from the goal.



(Q) How do we calculate the value (V) in this box?

(A) We calculate the value of this square first

For W: $V(s) = (0 + 0.9(1)) = 0.9$

For X: $V(s) = (0 + 0.9(0.9)) = 0.81$

For Y: $V(s) = (0 + 0.9(0.81)) = 0.73$

For Z: $V(s) = (0 + 0.9(0.81)) = 0.73$