Esercizio 2

b) Dato il seguente insieme di apprendimento

Esempio	*Target*	A_1 (4 *val.*)	A_2 (4 *val.*)	A_3 (3 *val.*)
1	$+$	v_1	v_1	v_2
2	$+$	v_3	v_3	v_1
3	$+$	v_4	v_1	v_3
4	$-$	v_2	v_1	v_3
5	$-$	v_2	v_2	v_2
6	$-$	v_2	v_3	v_2
7	$-$	v_3	v_4	v_1

mostrare come viene definito il classificatore bayesiano ingenuo usando una 7-stima con p uniforme

Soluzione

	P(| +)	P(| -)
A1 = v1	**11/40**	**7/44**
A1 = v2	**7/40**	**19/44**
A1 = v3	**11/40**	**11/44**
A1 = v4	**11/40**	**7/44**
A2 = v1	**15/40**	**11/44**
A2 = v2	**7/40**	**11/44**
A2 = v3	**11/40**	**11/44**
A2 = v4	**7/40**	**11/44**
A3 = v1	**10/30**	**10/33**
A3 = v2	**10/30**	**13/33**
A3 = v3	**10/30**	**10/33**
P(+) = 13/28 P(-) = 15/28		

c) Come viene classificato l'esempio (v_4 v_1 v_2) ? **Soluzione**

Ricordando che:

P(+) = 13/28 P(-) = 15/28

P(A1 = v4|+) = 11/40 P(A1 = v4|-) = 7/44

P(A2 = v1|+) = 15/40 P(A2 = v1|-) = 11/44

P(A3 = v2|+) = 10/30 P(A3 = v2|-) = 13/33

si ha:

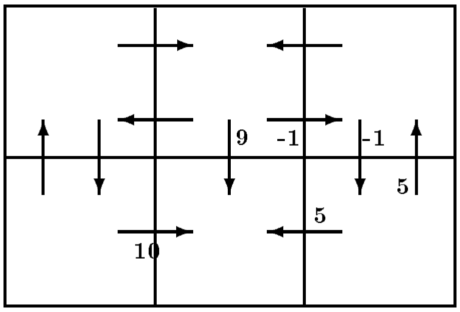
P(+)P(A1 = v4|+)P(A2 = v1|+)P(A3 = v2|+) = 0.015959

P(-) P(A1 = v4|-) P(A2 = v1|-)P(A3 = v2|-) = 0.0084

quindi v_{NB} = + e l'esempio e' classificato +

Esercizio 3

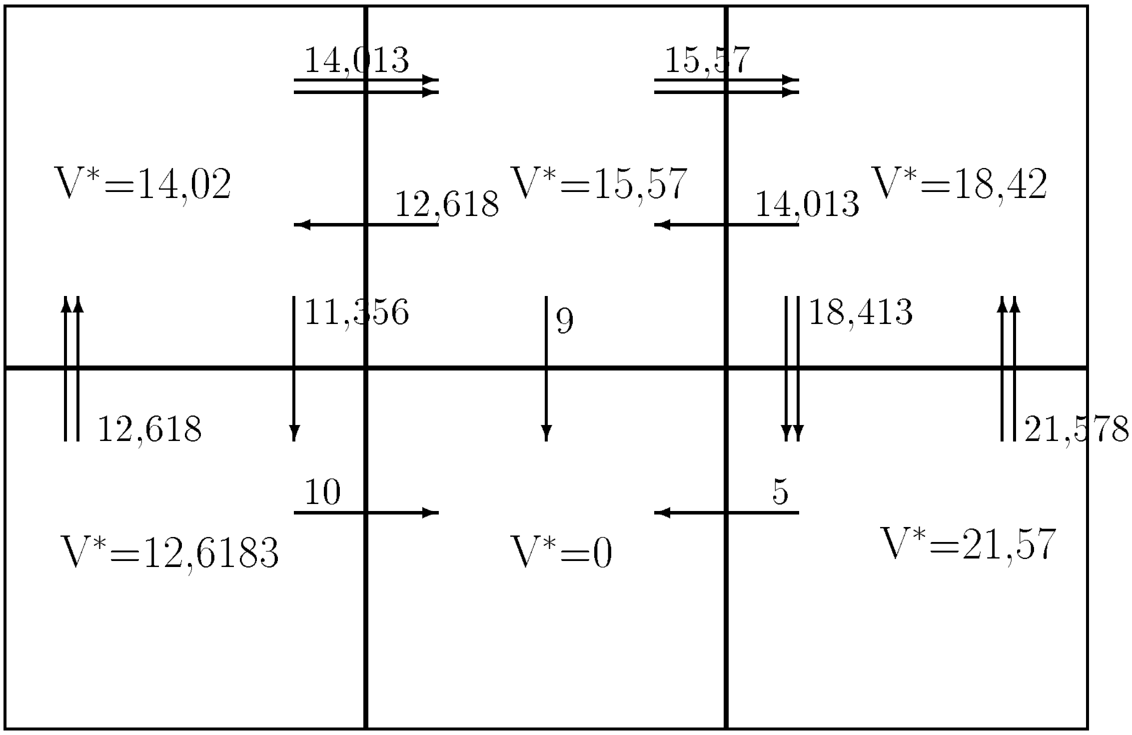
a) Si consideri il seguente ambiente con le corrispondenti azioni e ricompense



Calcolare i valori di V^* per ogni stato e $Q(s,a)$ per ogni coppia (s,a) avendo posto $\gamma = 0.9$.

Soluzione

Le doppie frecce individuano la politica ottima:



La soluzione si è ottenuta sfruttando le equazioni:

$$\Sigma_{i=0}^{\infty}\gamma^{2i} = \Sigma_{i=0}^{\infty}(\gamma^{2})^{i} = \frac{1}{1-\gamma^{2}}$$

$$\Sigma_{i=0}^{\infty}\gamma^{2i+1} = \gamma\,\Sigma_{i=0}^{\infty}\gamma^{2i}$$

b) Scoprire il valore minimo di γ per il quale la politica ottima definita in a) rimane valida, cioè quel valore γ^{min} tale che per $\gamma < \gamma^{min}$ la politica ottima in a) cambia.

Soluzione

Con il diminuire di γ, le ricompense ottenute per le azioni eseguite più in avanti con il tempo tenderanno ad essere ancora più scontate. Tale sconto può raggiungere un livello tale che una politica diversa, con ricompense immediate più vantaggiose, arrivi a prendere il sopravvento su quella corrente. In particolare si noti che lo stato più in basso a sinistra permette una azione, non presente nella politica ottima, con ricompensa immediata 10, così come lo stato al centro della parte superiore ne ha una con ricompensa immediata 9.

In generale tutte le azioni nella politica devono essere controllate. In questo caso, dai calcoli effettuati precedentemente si osserva direttamente che la prima azione che potrebbe entrare in gioco è quella più distante dal ciclo e con ricompensa immediata uguale a 10.

Affinché questa entri nella politica ottima deve essere vero che:

$$0 + \gamma V^*(\text{stato in alto a sinistra}) < 10$$

e quindi, poiché

$$V^*(\text{stato in alto a sinistra}) = 0 + \gamma(-1 + \gamma(-1 + \gamma\frac{5-\gamma}{1-\gamma^2}))$$

si ha che γ^{min} è quel valore massimo che soddisfa la seguente disuguaglianza

$$\gamma(\gamma(-1 + \gamma(-1 + \gamma\frac{5-\gamma}{1-\gamma^2}))) < 10$$