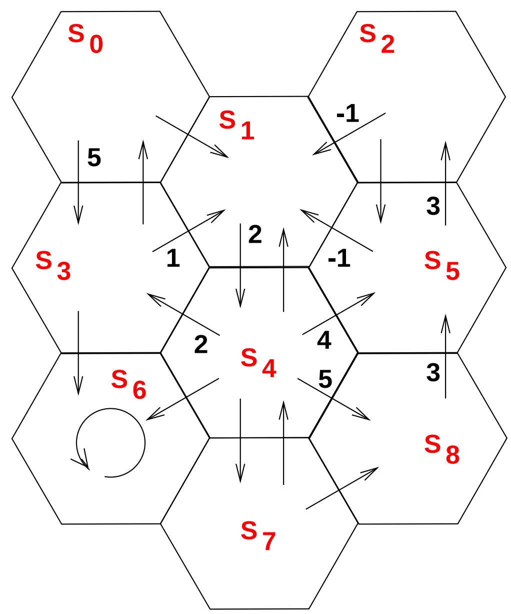
Esercizi con Soluzione del Corso di Sistemi per l'Elaborazione dell'Informazione

Terza Parte
(Reti Neurali e Apprendimento con Rinforzo)

Esercizio 1

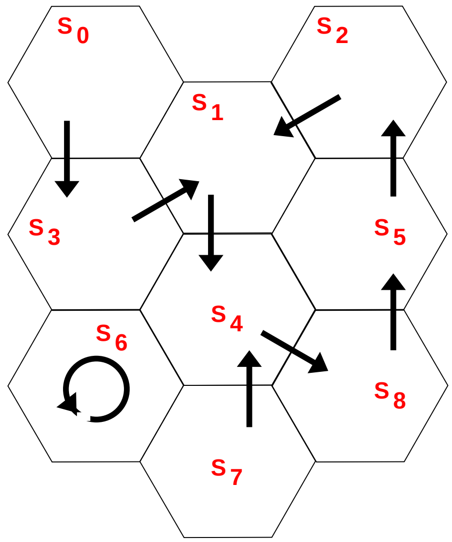
a1) Si consideri il seguente ambiente con le corrispondenti azioni e ricompense



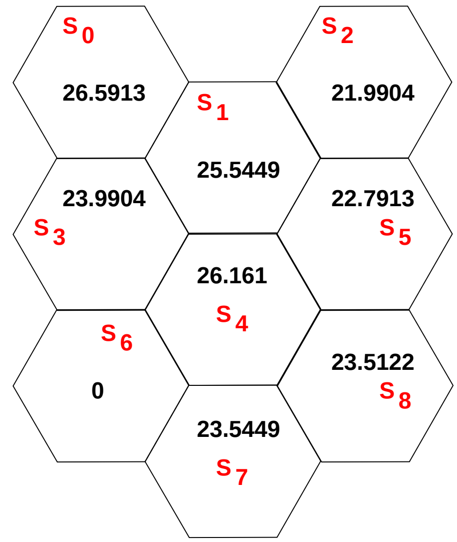
Calcolare i valori di V^* per ogni stato e $Q(s,a)$ per ogni coppia (s,a) avendo posto $\gamma = 0.9$.

Risposta:

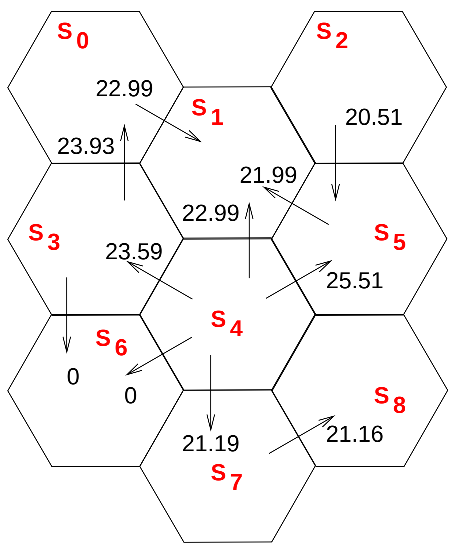
POLITICA OTTIMA



V*



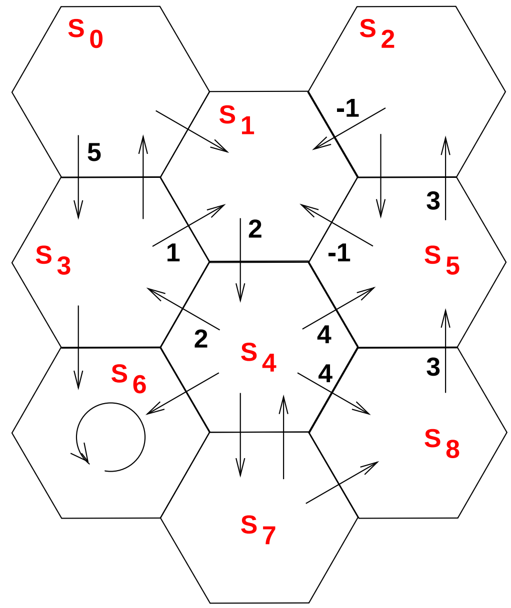
Q(s,a)



Le azioni della politica ottima hanno valore Q(s,a) uguale al valore V* dello stato di partenza. Si riportano i valori delle altre azioni.

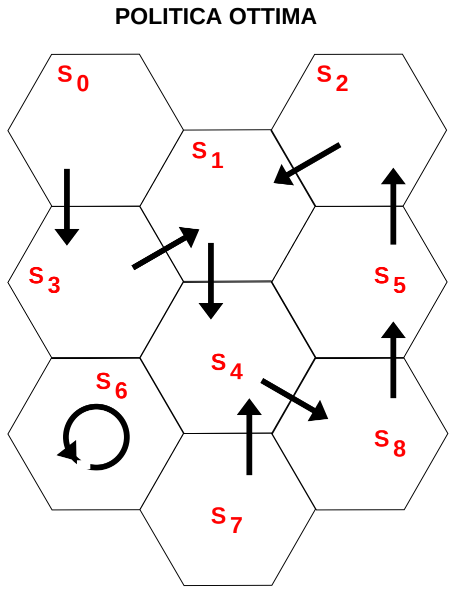
Esercizio 2

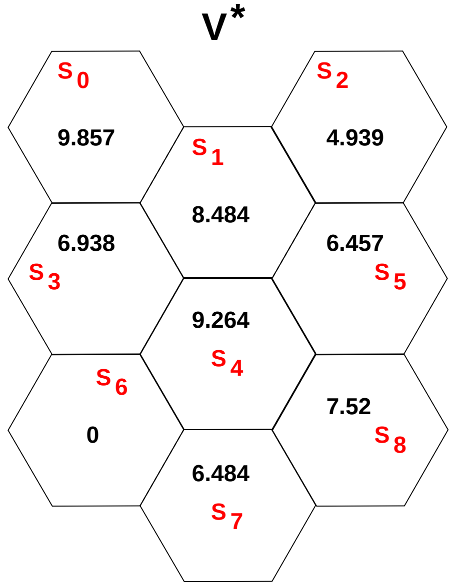
b2) Si consideri il seguente ambiente con le corrispondenti azioni e ricompense (le frecce senza numero aasociato indicano una ricompensa immediata di 0) Calcolare i valori

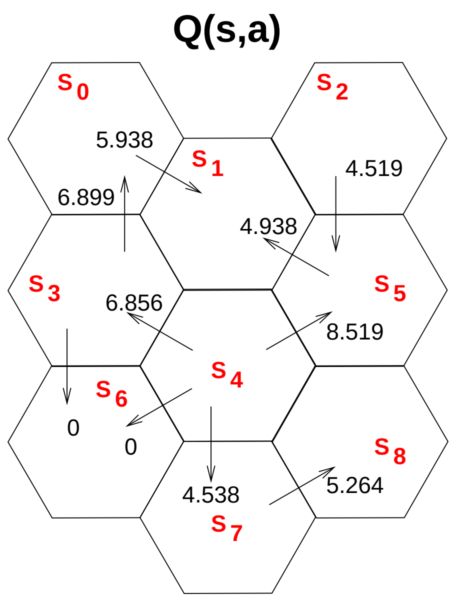


di V^* per ogni stato e $Q(s,a)$ per ogni coppia (s,a) avendo posto $\gamma = 0.7$.

Risposta:



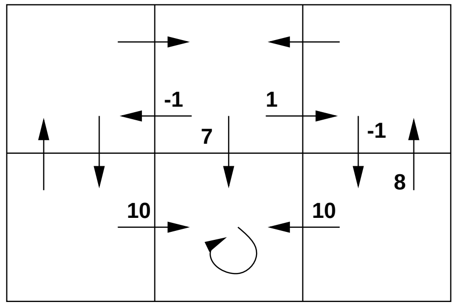




Le azioni della politica ottima hanno valore Q(s,a) uguale al valore V* dello stato di partenza. Si riportano i valori delle altre azioni.

Esercizio 3

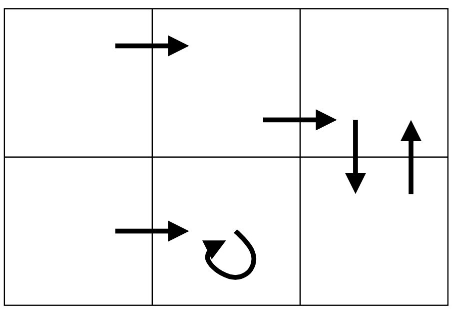
c2) Si consideri il seguente ambiente con le corrispondenti azioni e ricompense. Calcolare



i valori di V^* per ogni stato e $Q(s,a)$ per ogni coppia (s,a) avendo posto $\gamma = 0.8$;

Risposta:

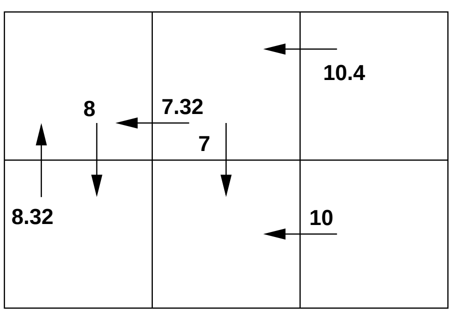
POLITICA OTTIMA



V*

10.4	13	15
10	0	20

Q(s,a)



Le azioni della politica ottima hanno valore Q(s,a) uguale al valore V* dello stato di partenza. Si riportano i valori delle altre azioni.

c3) Se con $(0,0)$ si indica lo stato in basso a sinistra, si mostri l'evoluzione dell'algoritmo Q-learning per la seguente sequenza (con reset dopo aver raggiunto lo stato $(1,0)$) di esplorazione:

$$(0,0), (0,1), (0,0), (1,0), reset, (2,1), (1,1), (2,1), (2,0), (1,0)$$

Risposta:

	→	←	↑	↓	↻
(0,0)		**no**		**no**	**no**
(1,0)	**no**	**no**	**no**	**no**	
(2,0)	**no**			**no**	**no**
(0,1)		**no**	**no**		**no**
(1,1)			**no**		**no**
(2,1)	**no**		**no**		**no**

Le caselle vuote contengono 0

s = (0,0) a = ↑

	→	←	↑	↓	↻
(0,0)		**no**		**no**	**no**
(1,0)	**no**	**no**	**no**	**no**	
(2,0)	**no**			**no**	**no**
(0,1)		**no**	**no**		**no**
(1,1)			**no**		**no**
(2,1)	**no**		**no**		**no**

Le caselle vuote contengono 0

s = (0,1) a = ↓

	→	←	↑	↓	↻
(0,0)	10	**no**		**no**	**no**
(1,0)	**no**	**no**	**no**	**no**	
(2,0)	**no**			**no**	**no**
(0,1)		**no**	**no**		**no**
(1,1)			**no**		**no**
(2,1)	**no**		**no**		**no**

Le caselle vuote contengono 0

s = (0,0) a = →

	→	←	↑	↓	↻
(0,0)	10	**no**		**no**	**no**
(1,0)	**no**	**no**	**no**	**no**	
(2,0)	**no**			**no**	**no**
(0,1)		**no**	**no**		**no**
(1,1)			**no**		**no**
(2,1)	**no**		**no**		**no**

Le caselle vuote contengono 0

s = (2,1) a = ←

	→	←	↑	↓	↻
(0,0)	10	**no**		**no**	**no**
(1,0)	**no**	**no**	**no**	**no**	
(2,0)	**no**			**no**	**no**
(0,1)		**no**	**no**		**no**
(1,1)	1		**no**		**no**
(2,1)	**no**		**no**		**no**

Le caselle vuote contengono 0

s = (1,1) a = →

	→	←	↑	↓	↻
(0,0)	10	**no**		**no**	**no**
(1,0)	**no**	**no**	**no**	**no**	
(2,0)	**no**			**no**	**no**
(0,1)		**no**	**no**		**no**
(1,1)	1		**no**		**no**
(2,1)	**no**		**no**	−1	**no**

Le caselle vuote contengono 0

s = (2,1) a = ↓

	→	←	↑	↓	↻
(0,0)	10	**no**		**no**	**no**
(1,0)	**no**	**no**	**no**	**no**	
(2,0)	**no**	10		**no**	**no**
(0,1)		**no**	**no**		**no**
(1,1)	1		**no**		**no**
(2,1)	**no**		**no**	−1	**no**

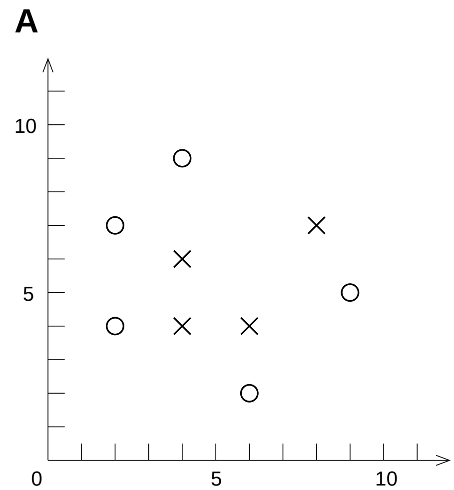
Le caselle vuote contengono 0

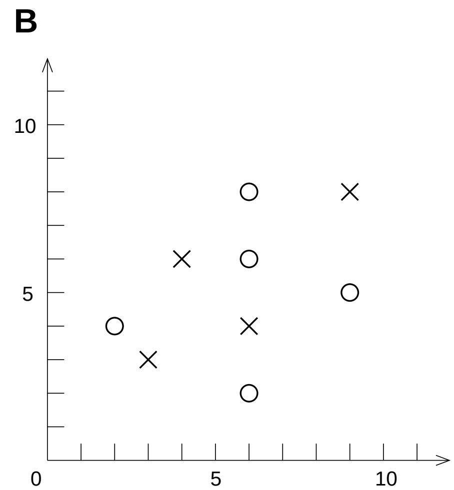
s = (2,0) a = ←

(0,1)	(1,1)	(2,1)
(0,0)	(1,0)	(2,0)

Esercizio 4

a) Dati i seguenti insiemi di apprendimento

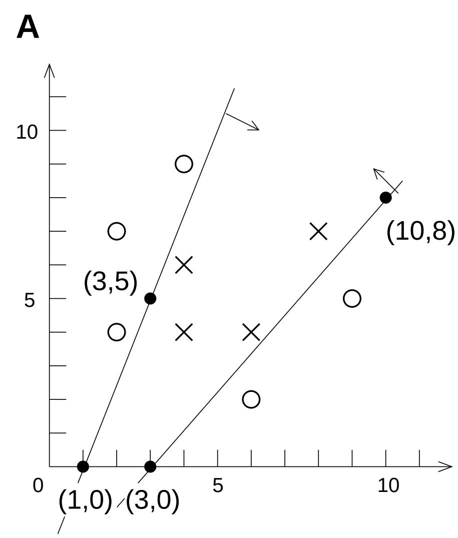


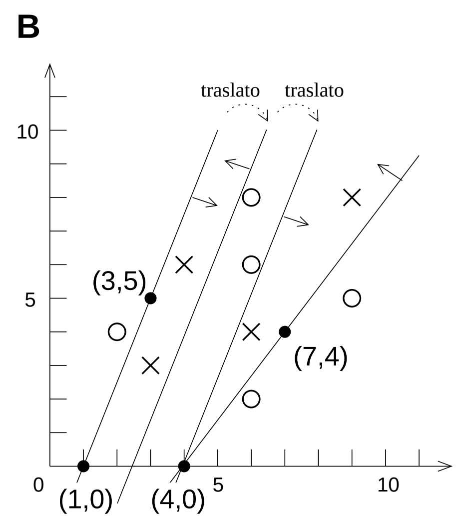


definire per ognuno di essi una rete neurale (con nodi a gradino) che li classifichi correttamente ($\times$ classificati positivamente).

Risposta:

I seguenti iperpiani classificano correttamente gli esempi dati:





caso A:

L'iperpiano più a sinistra passa per i punti $(1, 0)$ e $(3, 5)$, quindi, utilizzando la formula di una retta che passa per due punti (x_a, y_a), (x_b, y_b) data da $\frac{x-x_a}{x_b-x_a} = \frac{y-y_a}{y_b-y_a}$, abbiamo:

$$\frac{x_1 - 1}{3 - 1} = \frac{x_2 - 0}{5 - 0}$$

e quindi

$$x_1 - \frac{2}{5}x_2 - 1 = 0$$

da cui si deduce che $w_0 = -1,\ w_1 = 1,\ w_2 = -\frac{2}{5}$.

Per il secondo iperpiano abbiamo:

$$\frac{x_1 - 3}{10 - 3} = \frac{x_2 - 0}{8 - 0}$$

e quindi

$$x_1 - \frac{7}{8}x_2 - 3 = 0$$

da cui, considerando anche il verso del vettore tangente all'iperpiano, si deduce che $w_0 = 3,\ w_1 = -1,\ w_2 = \frac{7}{8}$.

caso B:

L'iperpiano più a sinistra è uguale al primo iperpiano del caso A. Il secondo e il terzo sono versioni traslate del primo iperpiano con eventualmente il verso cambiato di segno. Rispettivamente:

$$-x_1 + \frac{2}{5}x_2 + \frac{5}{2} = 0$$

da cui si deduce che $w_0 = \frac{5}{2},\ w_1 = -1,\ w_2 = \frac{2}{5}$,

$$x_1 - \frac{2}{5}x_2 - 4 = 0$$

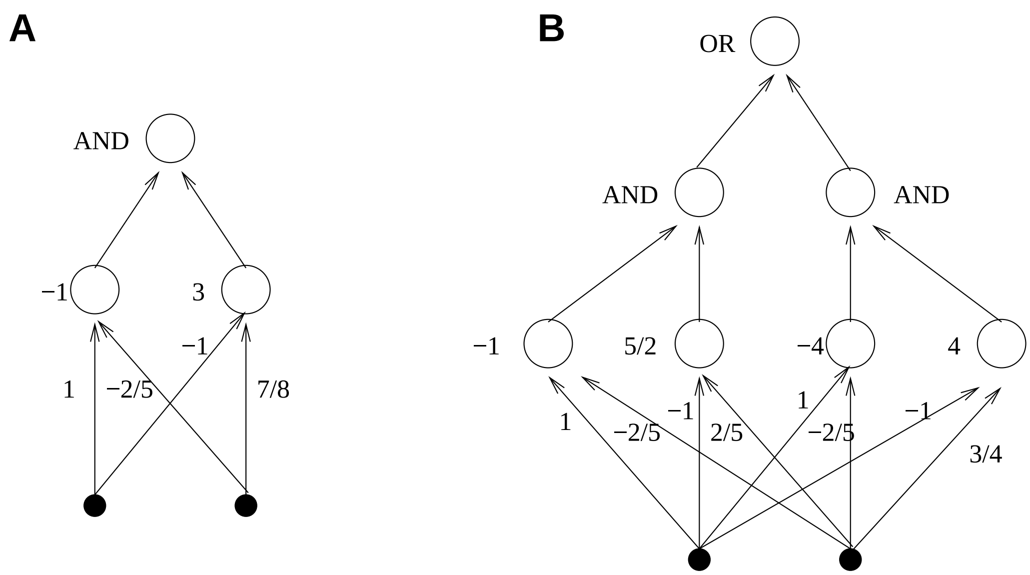
da cui si deduce che $w_0 = -4,\ w_1 = 1,\ w_2 = -\frac{2}{5}$.

Per l'ultimo iperpiano abbiamo:

$$-x_1 + \frac{3}{4}x_2 + 4 = 0$$

da cui si deduce che $w_0 = 4,\ w_1 = -1,\ w_2 = \frac{3}{4}$,

Infine, gli output dei vari iperpiani vanno opportunamente combinati tramite operatori AND e OR, come segue


A
AND
−1
3
1
−2/5
−1
7/8
B
OR
AND
AND
−1
5/2
−4
4
1
−2/5
−1
2/5
1
−2/5
−1
3/4