Actions:

* su
* giu
* dx
* sx

Stati 1 (cross)

* occupato (blocco, serpente)
* non occupato (libero, mela)

Stati 2

* vettore normalizzato della distanza con la mela (4 input)

Reward

* -100 ostacolo
* +10 mela
* Differenza della distanza di Manhattan tra lo step prima e quello dopo