Sprawozdanie - SIECI SAMOUCZĄCE SIĘ

W ćwiczeniu sailor, wykorzystałem metodę Montecarlo do rozwiązania zadania. Na samym początku tworzona jest losowa strategia –losowo wybierane są akcje spośród 4 akcji (przód, tył, lewo, prawo).  
Parametr **epsilon** odpowiada za wybór akcji. Określa prawdopodobieństwo z jakim akcja zostanie wybrana losowo zamiast według planowanej strategii. Parametr **alfa** jest współczynnikiem nagrody. Parametr **gamma** jest współczynnikiem dyskontowania.Następnie wykonywana jest funkcja environment zwracająca nagrodę i aktualny stan