LAPORAN TUGAS 3

REINFORCEMENT LEARNING: Q-LEARNING

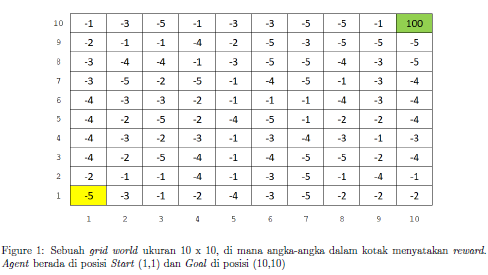
Nama : Adhyfa Fahmy Hidayat

Kelas : IF 39-01

NIM : 1301154127

## Analisis Masalah

Dalam tugas 3 ini, kita harus membuat sebuah sistem Q-Learning untuk menemukan optimum policy dari posisi start (kuning) sampai ke posisi goal (hijau). Agent hanya bisa melakukan 4 aksi: atas, bawah, kiri, kanan.



Yang harus dicari adalah si agent lewat mana saja sampai ke goals di posisi matriks (10,10)

## Desain

Algoritma Q-Learning secara umum adalah sebagai berikut:

* Tentukan parameter gamma
* Tentukan reward dalam matriks R
* Inisialisasi matriks Q ke 0
* Perulangan for:  
  per episode sampai goal

pilih random state sebagai current state (initial)

perulangan while:

sampai goal belum tercapai

* + Pilih action yang memungkinkan
  + Lalu lanjut bergerak ke action yang memungkinkan tersebut
  + Cari nilai Q yang paling maksimum berdasarkan seluruh action yang memungkinkan
  + Hitung nilai Q menggunakan rumus yang sudah ditentukan  
    
  + Tentukan next state sebagai current state

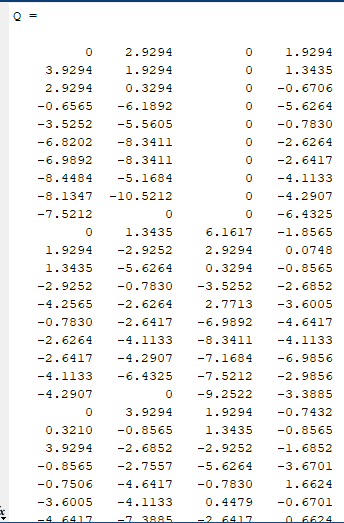
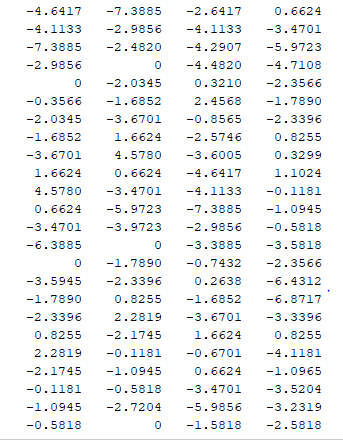
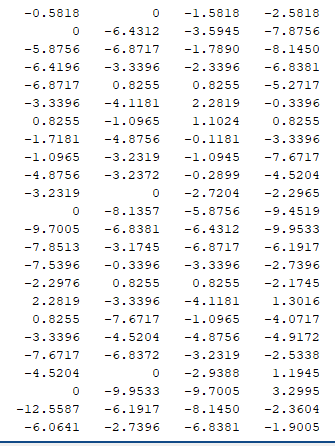
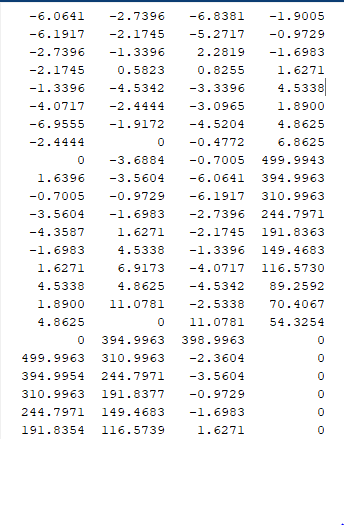
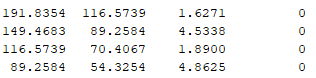
End while

End for

## Hasil Eksperimen

Pada file reward.xlsx berisi reward yang didapat dari table matriks DataTugasML3.txt pada setiap possible actionnya. Dapat dilihat bahwa goal berada di baris ke 91 dan initial state pada baris ke 10.

Berikut adalah Q akhir yang sudah dihitung dengan gamma = 0,8

Setelah mencapai current 91 yang dimana merupakan goal state maka perulangan berhenti. Setelah itu mendapatkan total reward dari initial state sampai goal state sebanyak:

