

## 1. Deskripsi Masalah

Diberikan sebuah data berisi 225 angka dengan format matriks  $15 \times 15$ , matriks tersebut adalah *reward* dari sebuah perjalanan, dimana perjalanan tersebut kita harus mendapatkan *reward* yang besar, maka berikut adalah algoritmanya :

- Set Qtable, dan set nilai tak hingga apabila tidak bisa melakukan pergerakan, misal di ujung atas, maka tidak bisa bergerak ke atas, Q table berisi 225 kolom (dari total  $15 \times 15$ ) dan 4 baris (4 arah gerak)

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
0	-inf	0	0	0	0	0	0	0	0														
1	0	0	0	0	0	0	0	0	0	0	0	0	0	-inf	0	0	0	0	0	0	0	0	
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
3	-inf	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-inf	0	0	0	0	0	0	

- Karena ada total 225 data, maka untuk mengambil nilai reward dibutuhkan sebuah array  $15 \times 15$  berisi urusan indeks dari state (poin x,y)0-224
- Set gamma =1 dan alpha =1
- Melakukan pencarian nilai *reward* maksimum dengan mengupdate Qtable
  - Melakukan random untuk menentukan step
  - Bila random = 0 maka state berkurang 15 (karena north berarti naik ke atas dan mengungarangi 15 langkah ke belakang karena state 0 berada di bawah)
  - Bila random= 1 maka state bertambah 1 (karena east, yang berarti bergerak ke kanan dan menambah state 1)
  - Bila random =2 maka state bertambah 15 (karena south berarti turun ke bawah dan menambah 15 langkah ke depan karena state 224 berada di atas)
  - Bila random =3 maka state berkurang 1 (karena west, bergerak ke kiri dan state berkurang 1 )
- Update Q table dengan persamaan :
  - $Q_{tab}[arah][state] = Q_{tab}[arah][state] + \alpha * (reward[rd][sacc] + (\gamma * \max(Q_{tab}[kolom\ state+1])) - Q_{tab}[arah][state]))$
- State = state+1

## 1. Analisis

Pada permasalahan kali ini terdapat sejumlah 225 data dimana berbentuk array dan dimulai pada kotak pertama di titik array dengan row = 14 column =0 dan berakhir di titik array dengan row =0 dan column =14 dengan reward 500. Kita diminta untuk mendapatkan *reward* sebesar besarnya.

Berikut adalah hasil runningnya :

Hasil Update Qtab dan hasil running

The screenshot shows a Q-table editor window titled "Qtab - NumPy array". The main area displays a 15x15 grid of numerical values. Row 0 contains mostly zeros with some NaN and -inf values. Row 1 has a notable entry of 500 at index (1, 14). Rows 2 and 3 are filled with -inf values. Below the grid is a scroll bar. At the bottom of the editor are buttons for "Format", "Resize", "Background color", "Save and Close", and "Close". To the right of the editor is a vertical log of actions:

```

East
state : 145
reward : -2.0

North
state : 160
reward : -4.0

North
state : 175
reward : -2.0

East
state : 176
reward : -2.0

North
state : 191
reward : -2.0

North
state : 206
reward : nan

North
state : 221
reward : nan

East
state : 222
reward : -3.0

East
state : 223
reward : 500.0

```

Didapatkan hasil running terbaik dengan hasil *reward* = 500, tidak dilakukan perulangan sebanyak n jumlah episode karena dengan algoritma ini, state akan menentukan sendiri jalan dan berhenti saat reward telah menyentuh nilai maksimum

Dalam program ini bisa terjadi looping forever, diduga terjadi karena hasil dari tabel bernilai nan, dan tidak terdeteksi nilai maksimumnya, double scalar error juga terjadi diduga karena mengalikan nilai dengan 0 atau nan, namun jika dilakukan running kembali,bisa menjadi normal lagi.