### Laporan Tugas Pemograman 3

#### Analisis Masalah

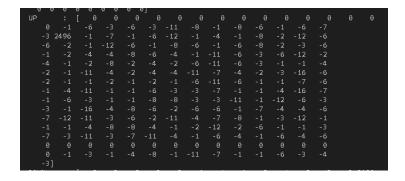
Bangunlah sebuah program Q-learning untuk menemukan optimum policy sehingga Agent yang berada di posisi Start (1,1) mampu menemukan Goal yang berada di posisi (15,15) dengan mendapatkan Total Reward maksimum pada grid world Gambar di bawah ini. Data pada gambar tersebut dapat dilihat di file DataTugas3ML2019.txt. Pada kasus ini, Agent hanya bisa melakukan empat aksi: N, E, S, dan W yang secara berurutan menyatakan North (ke atas), East (ke kanan), South (ke bawah), dan West (ke kiri). Anda boleh menggunakan skema apapun dalam mengimplementasikan sebuah episode.

# Penyelesaian

Yang pertama saya lakukan ada melakukan Pembuatan 2 Tabel Q dan R dengan insialasis table terdapat atas, kanan,kiri bawah. Setelah itu dilakukan fungsi pencrian jalan dengan batas matrix antara 0 sampai 14 kemudian dapat hasil sebuah nilai untuk memilih jalan dan ambil fungsi random terhadap pilihan jalan dan kita hituung nilai Q dengan pilihan yang ada diambil nilai terbaik kemudian dimasuk kedalam Q dan x dan y diubah sesuah dengan action terhitung satu episode dan reward merupakan nilai tambah dari currentstate yang pernah dilalui . akan berhenti ketika nilai current stat == 500

## Print Hasil:

## Hasil hitung Tabel Q



| R | IGHT    | :    | -1   | 9     | -1  | -2 | -4  | - 1 | -8  | - 1 | -12 | -6 | -4  | -2  | -6 | -2 |  |
|---|---------|------|------|-------|-----|----|-----|-----|-----|-----|-----|----|-----|-----|----|----|--|
|   | - 1     | -4   | . 0  | - 1   | -4  | -2 | - 1 | -2  | -4  | -2  |     | -2 | - 1 | -2  |    |    |  |
|   | -4      | -4   | -4   |       | -8  |    |     | -2  | -11 | -6  | -4  |    |     |     |    |    |  |
|   |         | -8   | -2 : | 1992  |     |    |     |     | -11 |     | - 1 | -4 |     | - 1 |    |    |  |
|   | -11     | -8   |      | -4    | -4  |    |     | -4  |     |     | -4  |    |     | - 1 |    |    |  |
|   | - 1     | -6   |      | -2    |     | -2 |     |     |     |     | -11 |    | - 1 | -8  |    |    |  |
|   | -11     |      |      | -2    |     | -7 |     |     |     |     | -12 |    | - 1 | -2  |    |    |  |
|   |         |      | - 1  | -4    | -12 |    |     |     |     | -12 | -6  |    |     | - 1 |    |    |  |
|   | -16     | -8   | -4   | -2    | -2  | -2 | -2  |     |     |     | -12 | -6 |     | -4  |    |    |  |
|   |         |      | -2   | -2    | -11 | -4 |     | -8  |     |     |     |    |     |     |    |    |  |
|   | - 16    | -8   | -4   | -8    |     | -2 | -4  | -6  |     | - 1 | - 1 |    |     | -11 |    |    |  |
|   |         |      |      |       | -16 |    |     | -8  |     |     | -8  | -2 |     | - 1 |    |    |  |
|   | -2      | -6   |      | -12   | -11 |    |     |     | -12 | -2  |     |    | -2  |     |    |    |  |
|   | -11     |      | -4   | -4    |     |    | -7  |     |     | -2  | -11 | -4 | -7  | -2  |    |    |  |
|   |         |      | -8   | -4    | -12 | -2 |     | -11 |     | -2  | -6  |    | - 1 | -7  |    |    |  |
|   |         |      |      |       |     |    |     |     |     |     |     |    |     |     |    |    |  |
|   | 1996]   |      |      |       |     |    |     |     |     |     |     |    |     |     |    |    |  |
| M | lax Re⊳ | ıard |      | : 350 | 9   |    |     |     |     |     |     |    |     |     |    |    |  |