

Proje ile paylaşılan dokumanda verilen input33.txt e (0-8 arası dokuz hücre) ait R matrisi aşağıdaki şekildedir (7 nolu hücre hedefimiz ise).

```
-1 -1 -1 0 -1 -1 -1 -1 -1
-1 -1 0 -1 0 -1 -1 -1 -1
-1 0 -1 -1 -1 0 -1 -1 -1
0 -1 -1 -1 0 -1 0 -1 -1
-1 0 -1 0 -1 -1 -1 -1 -1
-1 -1 0 -1 -1 -1 -1 -1 0
-1 -1 -1 0 -1 -1 -1 100 -1
-1 -1 -1 -1 -1 -1 0 100 -1
-1 -1 -1 -1 -1 0 -1 -1 -1
```

Algoritma her iterasyonda rastgele bir durumdan öğrenmeye başlar, Aksiyonlara göre durum değiştirir ve Q matrisini günceller (Başlangıçta Q matrisi sıfıra setlenmiştir). Her iterasyonda aşağıdaki işlemler tekrarlanır:

**i- rastgele bir durum sec**

1 nolu durum secilmiş olsun

**ii-secilen durumdan hedefe ulaşınca kadar Q matrisi aşağıdaki formüle göre güncellenir:**

$Q(\text{durum}, \text{aksiyon}) = R(\text{durum}, \text{aksiyon}) + \gamma \times \text{Max}\{Q(\text{sonraki-durumlar}, \text{tüm-aksiyonlar})\}$

1 nolu durumdan gidilebilecek yerler 2 ve 4 den birisi rasgele secilir (4 secilmiş olsun)

$Q(1, 4) = R(1, 4) + 0.8 \times \text{Max}\{Q(4, 1), Q(4, 3)\} \rightarrow R(1, 4)$  ve  $Q(4, 1), Q(4, 3)$  sıfır olduğundan  $Q(1, 4)$  sıfır olarak kalır.

4 nolu durumdan gidilebilecek yerler 1 ve 3 ten birisi rasgele secilir (1 secilmiş olsun)

$Q(4, 1) = R(4, 1) + 0.8 \times \text{Max}\{Q(1, 2), Q(1, 4)\} \rightarrow R(4, 1)$  ve  $Q(1, 2), Q(1, 4)$  sıfır olduğundan  $Q(4, 1)$  sıfır olarak kalır.

1 nolu durumdan gidilebilecek yerler 2 ve 4 den birisi rasgele secilir (4 secilmiş olsun)

$Q(1, 4) = R(1, 4) + 0.8 \times \text{Max}\{Q(4, 1), Q(4, 3)\} \rightarrow R(1, 4)$  ve  $Q(4, 1), Q(4, 3)$  sıfır olduğundan  $Q(1, 4)$  sıfır olarak kalır.

4 nolu durumdan gidilebilecek yerler 1 ve 3 ten birisi rasgele secilir (3 secilmiş olsun bu sefer)

$Q(4, 3) = R(4, 3) + 0.8 \times \text{Max}\{Q(3, 0), Q(3, 4), Q(3, 6)\} \rightarrow R(4, 3)$  ve  $Q(3, 0), Q(3, 4), Q(3, 6)$  sıfır olduğundan  $Q(4, 3)$  sıfır olarak kalır.

3 nolu durumdan gidilebilecek yerler 0, 4 ve 6 dan birisi rasgele secilir (6 secilmiş olsun)

$Q(3, 6) = R(3, 6) + 0.8 \times \text{Max}\{Q(6, 3), Q(6, 7)\} \rightarrow R(3, 6)$  ve  $Q(6, 3), Q(6, 7)$  sıfır olduğundan  $Q(3, 6)$  sıfır olarak kalır.

6 nolu durumdan gidilebilecek yerler 3 ve 7 den birisi rasgele secilir (7 secilmiş olsun)

$Q(6, 7) = R(6, 7) + 0.8 \times \text{Max}\{Q(7, 6), Q(7, 7)\} \rightarrow R(6, 7)$  yüz olduğundan  $Q(6, 7)$  100e setlenir. Hedefe ulaşıldığı için ilk iterasyon biter. İlk iterasyon sonucunda Q matrisi aşağıdaki gibidir.

```
0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 100 0
0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0
```

iii-Kullanıcı tarafından belirlenen iterasyon sonlanana kadar ilk 2 adım uygulanır.

Tum iterasyonlardan (3000 icin) sonra Q matrisi aşağıdaki gibidir.

```
0 0 0 316.8 0 0 0 0 0
0 0 160.8 0 252.8 0 0 0 0
0 201.6 0 0 0 128 0 0 0
252.8 0 0 0 252.8 0 396.8 0 0
0 201.6 0 316.8 0 0 0 0 0
0 0 160.8 0 0 0 0 0 102.4
0 0 0 316.8 0 0 0 496.8 0
0 0 0 0 0 396.8 496.8 0
0 0 0 0 0 128 0 0 0
```

Eger kullanıcı baslangıc hucresini 1 belirlemişse path su sekilde bulunur Q matrisi kullanılarak:

1 satırda max değeri (252.8) 4. sutundadır, o zaman bir sonraki hucre 4 tür.  
4. satırda max değeri (316.8) 3. sutundadır, o zaman bir sonraki hucre 3 tür.  
3. satırda max değeri (396.8) 6. sutundadır, o zaman bir sonraki hucre 6 dir.  
6. satırda max değeri (496.8) 7. sutundadır, o zaman bir sonraki hucre 7 dir.  
Hedef hucreye ulasıldığı icin path su sekildedir: 1 4 3 6 7

Benzer sekilde farklı baslangıc dugumleri icin hedef 7. dugum olacak sekilde path ler bulunabilir.