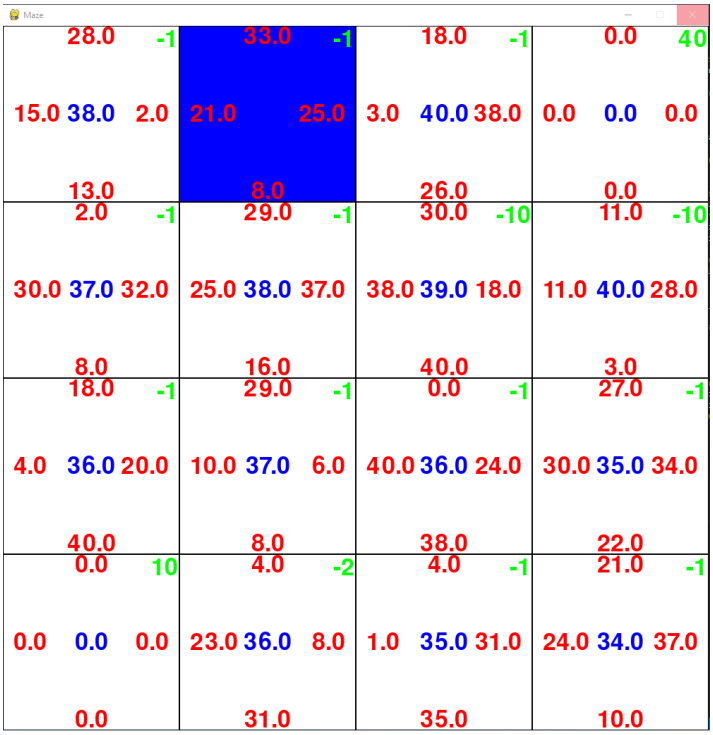
Toelichting antwoorden AS 2.2

Sam van Roosmalen
30-05-2023

Value Iteration



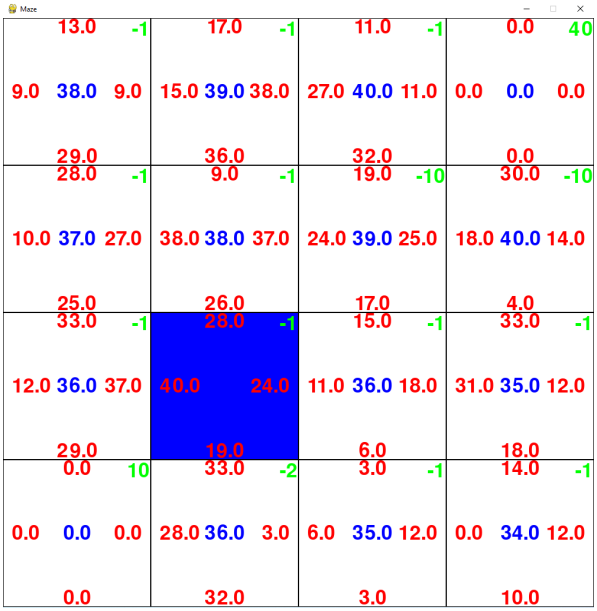
Policy table:

R = Right
L = Left
U = Up
D = Down

['R', 'R', 'R', 'D']
['R', 'U', 'U', 'U']
['R', 'U', 'L', 'L']
['D', 'U', 'U', 'L']

TD Learning

gamma = 1

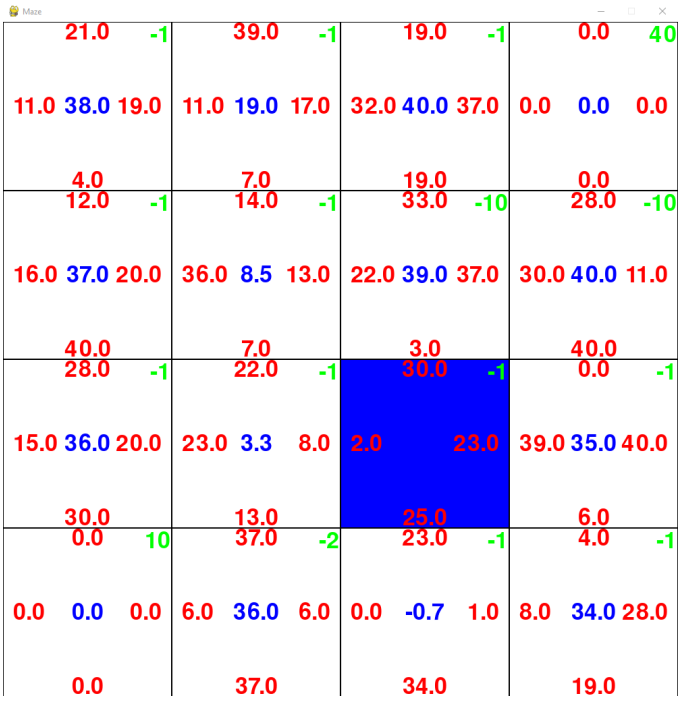


Policy table:

R = Right
L = Left
U = Up
D = Down

['R', 'R', 'R', 'D']
['R', 'U', 'U', 'U']
['R', 'U', 'L', 'L']
['D', 'U', 'U', 'L']

gamma = 0.5

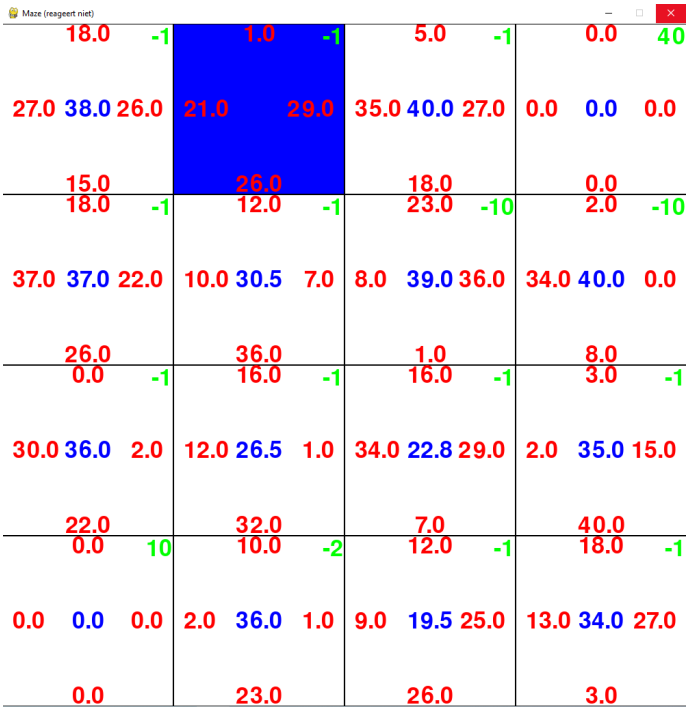


Policy table:

R = Right
L = Left
U = Up
D = Down

['R', 'R', 'R', 'D']
['R', 'U', 'U', 'U']
['R', 'U', 'L', 'L']
['D', 'U', 'U', 'L']

gamma = 0.9



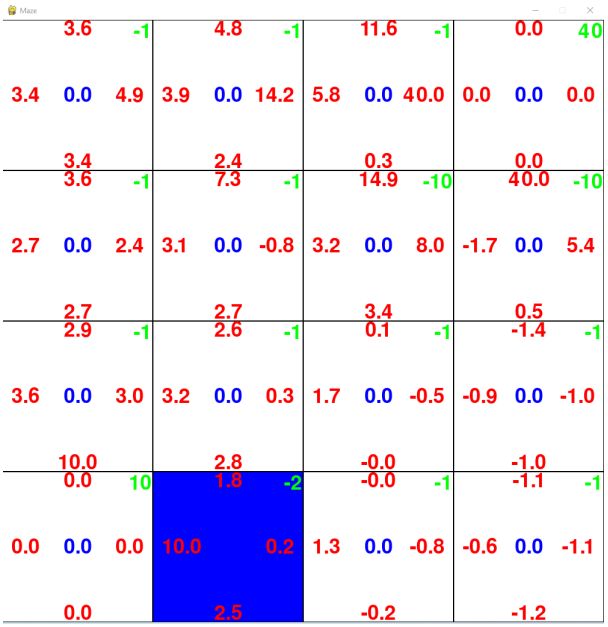
Policy table:

R = Right
L = Left
U = Up
D = Down

['R', 'R', 'R', 'D']
['R', 'U', 'U', 'U']
['R', 'U', 'L', 'L']
['D', 'U', 'U', 'L']

Sarsa Control

gamma = 1

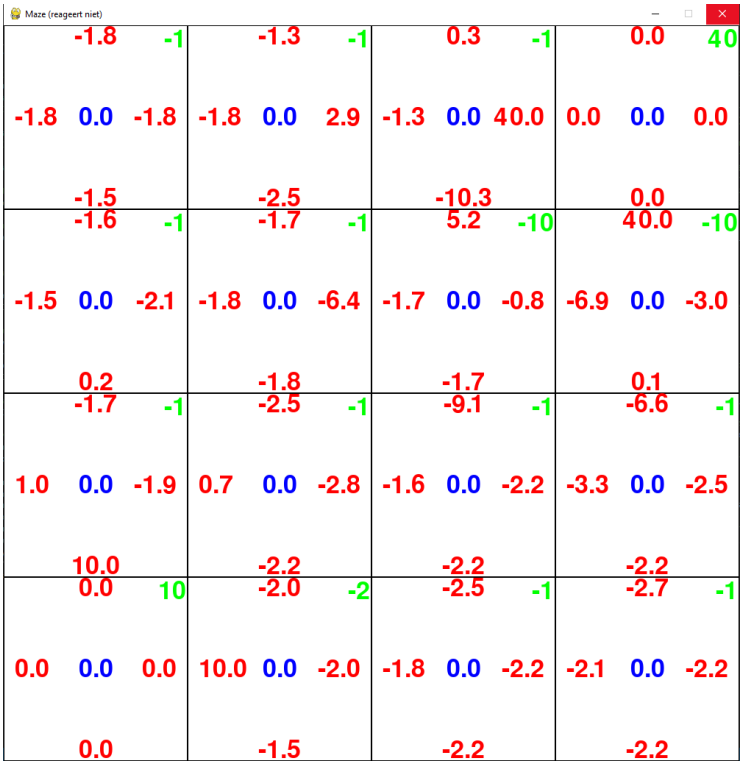


Policy table:

R = Right
L = Left
U = Up
D = Down

['R', 'R', 'R', '-']
['U', 'U', 'U', 'U']
['D', 'L', 'L', 'L']
['-', 'L', 'L', 'L']

gamma = 0.5

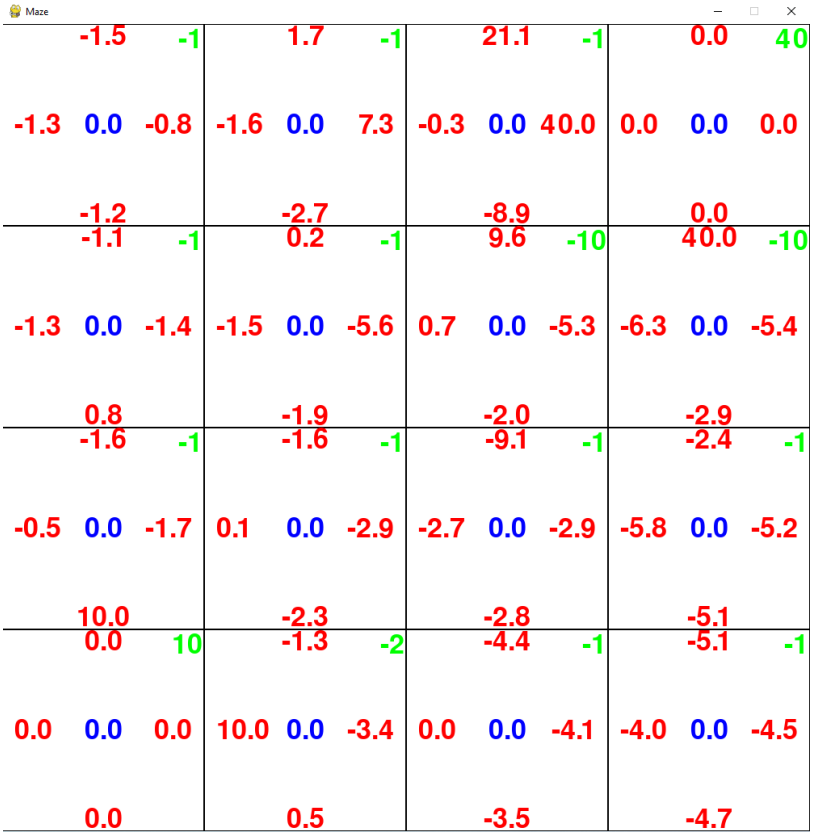


Policy table:

R = Right
L = Left
U = Up
D = Down

['R', 'R', 'R', '-']
['U', 'U', 'U', 'U']
['D', 'L', 'L', 'L']
['-', 'L', 'L', 'L']

gamma = 0.9



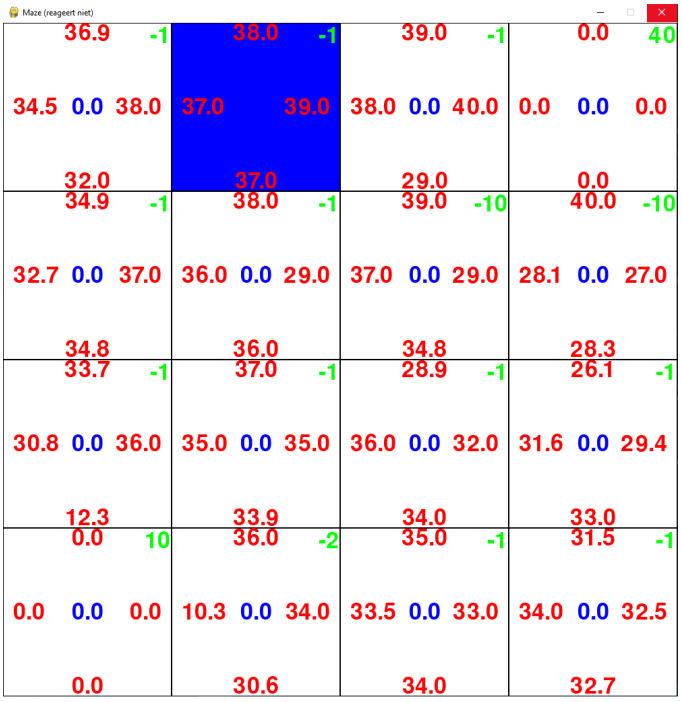
Policy table:

R = Right
L = Left
U = Up
D = Down

['R', 'R', 'R', '-']
['D', 'U', 'U', 'U']
['D', 'L', 'L', 'U']
['-', 'L', 'L', 'L']

Q Learning

gamma = 1

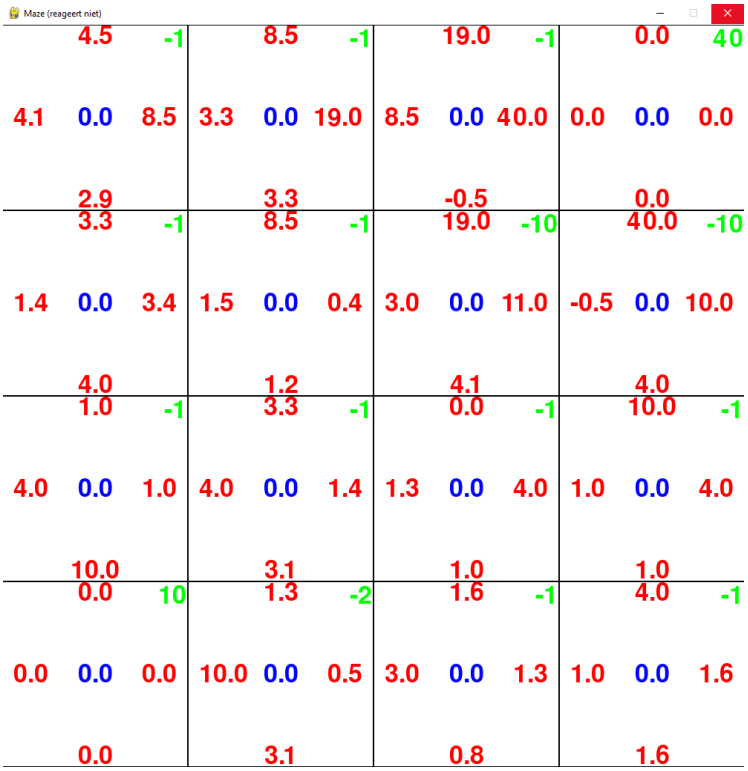


Policy table:

R = Right
L = Left
U = Up
D = Down

['R', 'R', 'R', '-']
['R', 'U', 'U', 'U']
['R', 'U', 'L', 'D']
['-', 'U', 'U', 'L']

gamma = 0.5

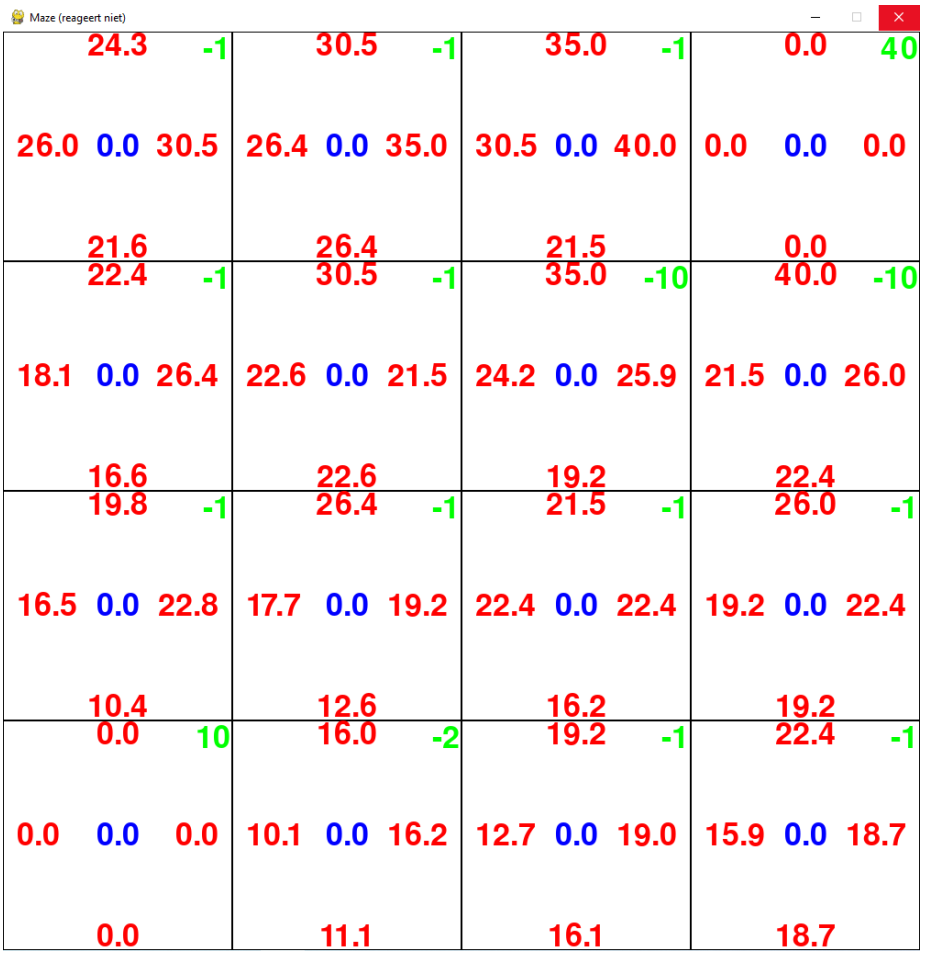


Policy table:

R = Right
L = Left
U = Up
D = Down

['R', 'R', 'R', '-']
['D', 'U', 'U', 'U']
['D', 'L', 'R', 'U']
['-', 'L', 'L', 'U']

gamma = 0.9



Policy table:

R = Right
L = Left
U = Up
D = Down

['R', 'R', 'R', '-']
['R', 'U', 'U', 'U']
['R', 'U', 'R', 'U']
['-', 'R', 'U', 'U']