B # -\*- coding: utf-8 -\*-

"""

Created on Tue Jul 25 15:42:59 2023

@author: HP

"""

#TD PREDICTION

import gymnasium as gym

import pandas as pd

#create the envt

env = gym.make("FrozenLake-v1", render\_mode = "human")

env.reset()

env.render()

def random\_policy():

return env.action\_space.sample()

V = {}

for s in range(env.observation\_space.n):

V[s] = 0.0

alpha = 0.85

gamma = 0.90

num\_eps = 50

num\_steps = 10

for i in range(num\_eps):

s = env.reset()

s = s[0]

for t in range(num\_steps):

a = random\_policy()

s\_, r, done, \_, \_ = env.step(a)

V[s] += alpha \* (r + gamma \* V[s\_] - V[s])

s = s\_

if done:

break

df = pd.DataFrame(list(V.items()), columns = ['state', 'value'])

print(df)

#Sample Output after 5000 episodes

# state value

#0 0 0.000409

#1 1 0.000005

#2 2 0.007377

#3 3 0.000351

#4 4 0.003835

#5 5 0.000000

#6 6 0.004941

#7 7 0.000000

#8 8 0.000642

#9 9 0.014566

#10 10 0.112160

#11 11 0.000000

#12 12 0.000000

#13 13 0.059375

#14 14 0.529872

#15 15 0.000000