

---

```
% Achyuth Nandikotkur
% V00975928
% ECE-559B
% November 8th, 2021

% Question 3

clc;
clear;

w = [0; 0];
alpha = 0.05;
steps = 4000;
statesWithPolicies = cell(5,5);
states = 0:24;
gamma = 0.8

for i=1:numel(statesWithPolicies)
    statesWithPolicies{i} = {states(i), 0};
end

wx = [0];
wy = [0];

for outerloop = 1 : 1
    episode = generate_episode(4000);
    for step = 1:length(episode)-1
        r = rem(episode{step}.state, 5);
        q = (episode{step}.state-r)/5;

        % current state x-cordinate & y-coordinate
        x_coordinate = q;
        y_coordinate = r + 1;

        % next state x-coordinate & y-coordinate
        r = rem(episode{step+1}.state, 5);
        q = ((episode{step+1}.state)-r)/5;
        x_coordinate_next = q;
        y_coordinate_next = r + 1;

        target_value = transpose([wx(end); wy(end)]) *
[x_coordinate_next; y_coordinate_next];
        current_value = transpose([wx(end); wy(end)]) * [x_coordinate;
y_coordinate];

        wnew = [wx(end); wy(end)] + (alpha * (episode{step}.reward
+ (gamma * target_value) - current_value) * [x_coordinate;
y_coordinate]);
        wx =[wx; wnew(1)];
        wy =[wy; wnew(2)];
    end
end
```

---

---

```

end

for i = 1 : 5
    for j = 1 : 5
        statesWithPolicies{i, j}{2} = [wx(end) wy(end)] * [i; j];
    end
end

figure(1)
plot(1:length(wx), wx, 1:length(wy), wy);
xlabel('Episodes')
ylabel('State values')
legend({'w-x', 'w-y'}, 'Location', 'southwest')

function episode = generate_episode(steps)
    % using random policy

    % 4 step sequence
    sequence = cell(1, steps);

    % Since SARSA is single step, generating a single step episode
    for k1 = 1:steps
        sequence{k1}.state = 0;
        sequence{k1}.action = 0;
        sequence{k1}.reward = 0;
    end

    sequence{1}.state = randi(25) - 1;

    for i = 1:steps
        action = action_string(randi(4));
        sequence{i}.action = action;
        obj = calc_next_state(sequence{i}.state, action);
        sequence{i+1}.state = obj(1);
        sequence{i}.reward = obj(2);
    end

    episode = sequence;
end

function state_reward = calc_next_state(state, action)
    reward = 0;
    current_state = state;
    if(1 == current_state)
        next_state = 21;
        reward = 10;
    elseif(3 == current_state)
        next_state = 13;
        reward = 5;
    elseif(2 == current_state)
        reward = 0;
        if(strcmp(action, 'right'))

```

---

---

```

        next_state = state+1;
elseif(strcmp(action, 'left'))
    next_state = state-1;
elseif(strcmp(action, 'up'))
    next_state = state;
    reward = -1;
else
    next_state = state+5;
end
% right side
elseif(any([9, 14, 19] == current_state))
    reward = 0;
    if(strcmp(action, 'right'))
        next_state = state;
        reward = -1;
    elseif(strcmp(action, 'left'))
        next_state = state-1;
    elseif(strcmp(action, 'up'))
        next_state = state - 5;
    else
        next_state = state+5;
    end
end
% bottom side
elseif(any([21, 22, 23] == current_state))
    reward = 0;
    if(strcmp(action, 'right'))
        next_state = state + 1;
    elseif(strcmp(action, 'left'))
        next_state = state-1;
    elseif(strcmp(action, 'up'))
        next_state = state-5;
    else
        next_state = state;
        reward = -1;
    end
end
% left side
elseif(any([5, 10, 15] == current_state))
    reward = 0;
    if(strcmp(action, 'right'))
        next_state = state + 1;
    elseif(strcmp(action, 'left'))
        next_state = state;
        reward = -1;
    elseif(strcmp(action, 'up'))
        next_state = state-5;
    else
        next_state = state+5;
    end
end
% corners
elseif(0 == current_state)
    reward = 0;
    if(strcmp(action, 'right'))
        next_state = state + 1;
    elseif(strcmp(action, 'left'))

```

---

---

```

        next_state = state;
        reward = -1;
elseif(strcmp(action, 'up'))
    next_state = state;
    reward = -1;
else
    next_state = state+5;
end
elseif(4 == current_state)
    reward = 0;
    if(strcmp(action, 'right'))
        next_state = state;
        reward = -1;
    elseif(strcmp(action, 'left'))
        next_state = state - 1;
    elseif(strcmp(action, 'up'))
        next_state = state;
        reward = -1;
    else
        next_state = state + 5;
    end
elseif(24 == current_state)
    reward = 0;
    if(strcmp(action, 'right'))
        next_state = state;
        reward = -1;
    elseif(strcmp(action, 'left'))
        next_state = state - 1;
    elseif(strcmp(action, 'up'))
        next_state = state - 5;
    else
        next_state = state;
        reward = -1;
    end
elseif(20 == current_state)
    reward = 0;
    if(strcmp(action, 'right'))
        next_state = state + 1;
    elseif(strcmp(action, 'left'))
        next_state = state;
        reward = -1;
    elseif(strcmp(action, 'up'))
        next_state = state - 5;
    else
        next_state = state;
        reward = -1;
    end
% All other cases
else
    reward = 0;
    if(strcmp(action, 'right'))
        next_state = state + 1;
    elseif(strcmp(action, 'left'))
        next_state = state - 1;

```

---

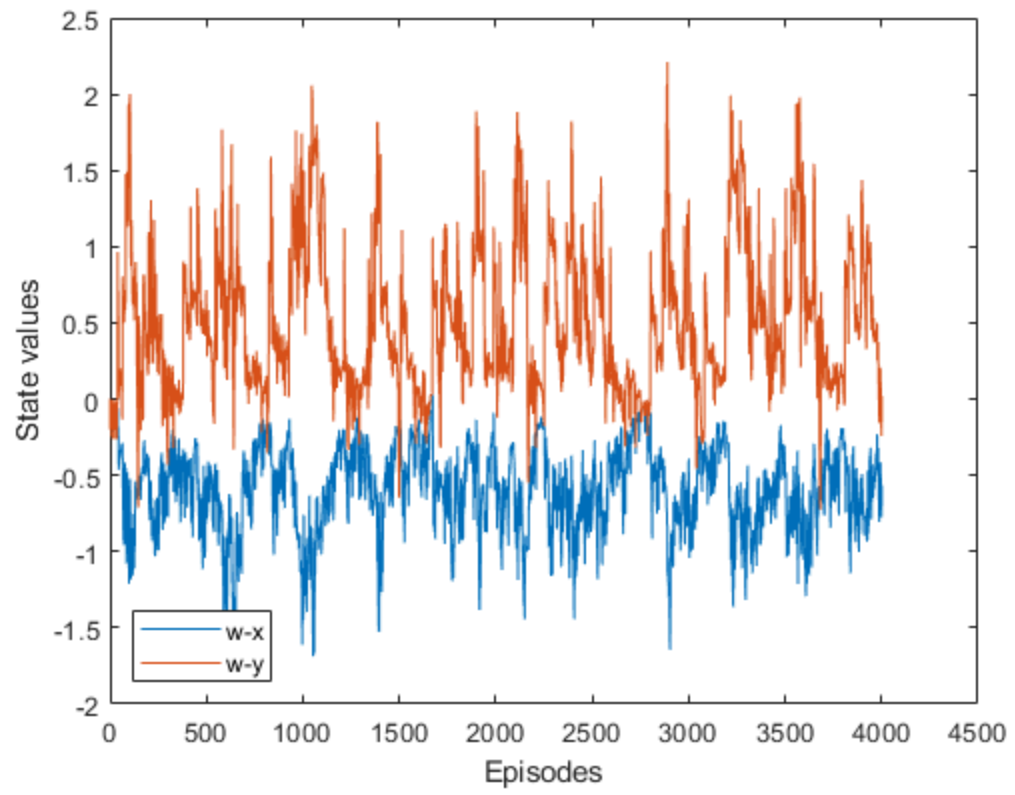
---

```
        elseif(strcmp(action, 'up'))
            next_state = state - 5;
        else
            next_state = state + 5;
        end
    end
    state_reward = [next_state, reward];
end

function action = action_string(num)
    res = '';
    if(num == 1)
        res = 'right';
    elseif(num == 2)
        res = 'left';
    elseif(num == 3)
        res = 'up';
    else
        res = 'down';
    end
    action = res;
end

gamma =

    0.8000
```



*Published with MATLAB® R2021a*