

```

function [Q,firstSARewCnt,firstSARewSum] = mcEstQ(stateseen,act_taken,rew,
firstSARewCnt,firstSARewSum,Q, maxNPii,maxNPjj,maxNVii,maxNVjj)
% MCESTQ – monte carlo estimate of the action value function for the race track example
%
% Written by:
% --
% John L. Weatherwax          2007-12-07
%
% email: wax@alum.mit.edu
%
% Please send comments and especially bug reports to the
% above email address.
%
%-----

% accumulate these values used in computing statistics on this action value function
Q^{pi}:
% Note 1) in this state transition there is no difference between first occurrence and
each occurrence ...
for si=1:size(stateseen,1),
    pii=stateseen(si,1); pjj=stateseen(si,2); vii=stateseen(si,3); vjj=stateseen(si,4);
    staInd = sub2ind( [maxNPii,maxNPjj,maxNVii,maxNVjj], pii,pjj,vii+1,vjj+1 );
    actInd = act_taken(si);
    firstSARewCnt(staInd,actInd) = firstSARewCnt(staInd,actInd)+1;
    firstSARewSum(staInd,actInd) = firstSARewSum(staInd,actInd)+rew;
    %Q(staInd,actInd)
    =
firstSARewSum(staInd,actInd)/firstSARewCnt(staInd,actInd); % <-take the direct average
    Q(staInd,actInd)
    = Q(staInd,actInd) + (1/firstSARewCnt(staInd,actInd))*(rew
- Q(staInd,actInd)); % <-use incremental averaging
    %Q(staInd,actInd)
    = Q(staInd,actInd) + alpha*(rew - Q(staInd,actInd)); %
<-take a geometric average
end

```