

```

%
% 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.
%
%-----

%close all;

% the number of experiments:
N_EXPS = 1000;

% the number of learning episodes (for complete learning this would need to be much
larger):
nEpisodes = 10;

% the number of random walk states (not including the two terminal states):
%n_rw_states = 17;
n_rw_states = 19;

% the true solution:
% (when zero is the reward for going off the left end):
Vtruth = (0:(n_rw_states+1))/(n_rw_states+1); Vtruth(1)=[]; Vtruth(end)=[];
% (when -1 is the reward for going off the left end):
Vtruth = -1 + 2*(0:(n_rw_states+1))/(n_rw_states+1); Vtruth(1)=[]; Vtruth(end)=[];

% a vector of learning parameters
%alphaV = linspace( 0, 0.3, 20 );
alphaV = linspace( 0, 0.3, 100 );
%alphaV = linspace( 0, 0.3, 10 );

% a vector of TD(n) lengths
nVect = [ 1 2 3 4 6 8 15 30 60 100 300 1000 ];

figure; hold on; results=zeros(length(alphaV),length(nVect));
for ni=1:length(nVect),
    n=nVect(ni);
    fprintf('working on n=%d...\n',n);
    tra = zeros(1,length(alphaV));
    for ai=1:length(alphaV),
        a = alphaV(ai);
        fprintf('working on alpha=%f...\n',a);
        tr = zeros(1,N_EXPS);
        for ei=1:N_EXPS,
            V = rw_offline_ntd_learn( n, a, n_rw_states, nEpisodes );
            tr(ei) = sqrt( mean( (V-Vtruth).^2 ) );
        end
        tra(ai) = mean( tr );
    end
    results(:,ni) = tra(:);
end

phs=plot( alphaV, results );
xlabel( 'alpha' ); ylabel( ['average rms error from ',num2str(nEpisodes), ' episodes'] );

```

```
nVectCL={};
for ti=1:length(nVect),
    nVectCL{end+1} = ['TD(n=',num2str(nVect(ti)),')'];
end
legend(phs,nVectCL,'Location','Best');
axis( [ 0, alphaV(end), 0.2, 0.55 ] );

saveas(gcf,['./rw_offline_ntd_learn_',num2str(N_EXPS)], 'png');

return;

% a simple point case to run:
n = 3;
alpha = 0.1;
rw_offline_ntd_learn
```