

**Algorithm** *NN-TD*(*G*, *J*)

(\* Estimates the value function *J* from the trajectories \*)

1. Let  $Y_n$  be the set of trajectories starting at  $x_{start}$  till leaf nodes in *G*.
2. **for** each trajectory  $(s_o, a_o, r_1, s_1, a_1, r_2, s_2, \dots)$  in  $Y_n$
3.       **for** each pair  $(s_i, a_i, r_{i+1}, s_{i+1})$
4.                $J(s_i) = (1 - \alpha)J(s_i) + \alpha(r_i + \gamma J(s_{i+1}))$
5.   return *J*
6.   End of Algorithm