Computational Intelligence in Games Summer 2017

Prof. Dr. Sanaz Mostaghim, Alexander Dockhorn

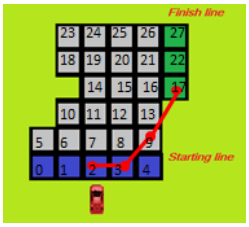
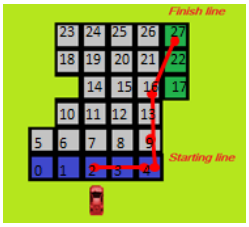
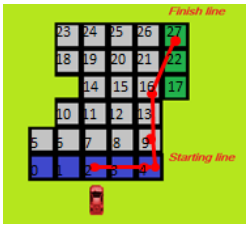
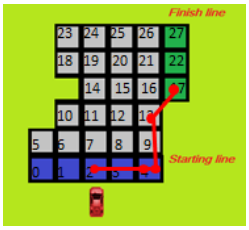
3. Exercise Sheet

Assignment 14 Monte Carlo method

In Monte Carlo method, if we start with a deterministic π, some/many $(s,\ a)$-pairs will never be visited! How can we make sure that (almost) all pairs are visited?

Assignment 15 Race track (MC)

- States: grid squares, velocity horizontal and vertical
- Rewards: -1 on track, -5 off track
- Only the right turns allowed
- Actions: +1, -1, 0 to velocity
- $0 <$ Velocity < 2 in each direction
- Stochastic: 50% of the time it moves 1 extra square up or right
- Goal: reach the finish line as fast as possible without leaving the track
- No discounting ($\gamma = 1$)
- Return for each state is the negative number of steps to go from that state
- V(s): predicted negative number of steps

			
You start from cell 2 with no velocity	After one move, west wind brings you to cell 4 (instead of 3)	You stay in cell 4	A wind from south moves you to cell 13 (instead of 9)
Estimated return of -3	Estimated return of $-4 \rightarrow 5$ steps in total	Estimated return of $-3 \rightarrow 5$ steps in total	Estimated return of -1 $\rightarrow$4 steps in total
You add +1 to horizontal velocity.	You add -1 to horizontal velocity.	You add +1 to vertical velocity.	You add +1 to horizontal velocity.
	Reward -1	Reward -1	Reward -1
State $= (2, 0, 0)$	State $= (4, 1, 0)$	State $= (4, 0, 0)$	State $= (13, 0, 1)$ Final State $= (17, 1, 1)$

Computational Intelligence in Games Summer 2017

Prof. Dr. Sanaz Mostaghim, Alexander Dockhorn

a) Complete the table below for the Race track example:

State $s(cell, h, v)$	Rewards so far	G_t	V(s)
$(2, 0, 0)$	0	-3	-3
$(4, 1, 0)$			
$(4, 0, 0)$			
$(13, 0, 1)$			
$(17, 1, 1)$			

b) Compute an iteration of Monte Carlo with $\alpha = 0.5$ for the Race track example

Iteration	$K = 0$	$K = 1,\ (\alpha = 0.5)$
$G_t(2, 0, 0)$		
$V(2, 0, 0)$		
$G_t(4, 1, 0)$		
$V(4, 1, 0)$		
$G_t(4, 0, 0)$		
$V(4, 0, 0)$		
$G_t(13, 0, 1)$		
$V(13, 0, 1)$		

Assignment 16 Race track (TD)

Compute an iteration of $TD(0)$ with $\alpha = 0.5$ for the Race track example

Iteration	$K = 0$	$K = 1,\ (\alpha = 0.5)$	Error δ
$V(2, 0, 0), \quad R_1 = -1$			
$V(4, 1, 0), \quad R_2 = -1$			
$V(4, 0, 0), \quad R_3 = -1$			
$V(13, 0, 1), \quad R_4 = -1$			