Stochastic Control and Optimization Course (EL2800)

Computer Lab 1 – The Great Escape

November 22, 2016

Department of Automatic Control
School of Electrical Engineering
KTH The Royal Institute of Technology

Solve at least Problem 1. Solving Problem 2 would get you 3 extra points at the exam (out of 50 points). Please send your answers and your code by email to magur@kth.se and mstms@kth.se before Friday December 2, 5PM. Good luck!

Problem 1: The Maze and the Random Minotaur

Consider the maze of Figure 1. You enter the maze in A and at the same time, the minotaur enters in B. The minotaur follows a random walk while staying within the limits of the maze. The minotaur's walk goes through walls (which obviously you cannot do). At each step, you observe the position of the minotaur, and decide on a one-step move (up, down, right or left) or not to move. Your objective is to identify the strategy maximising the probability to exit the maze before time T.

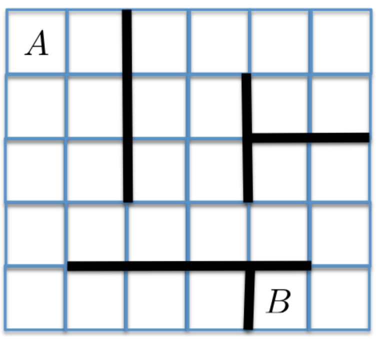

Figure 1: The maze

1. Formulate the problem as an MDP.
2. Solve the problem, and display the value function and an optimal policy for $T = 15$. Plot the maximal probability of exiting the maze as a function of T.
3. Assume now that your life is geometrically distributed with mean 30. Modify the problem so as to derive a policy minimising the expected time to exit the maze.

Problem 2: Robbing Banks

At time 0, you are robbing bank 1 (see Figure 2), and the police gets alerted and starts chasing you from the point PS (Police Station). You observe where the police is, and decide in each step either to move up, left, right, down or to stay where you are. Each time you are at a bank, and that the police is not there, you collect a reward of 10$. If the police catches you, you loose 50$, and the game is reinitialised (you go back to bank 1, and the police goes back to the PS).The rewards are discounted at rate $\lambda \in (0, 1)$. The police always chases you, but moves randomly in your direction. More precisely, assume that the police and you are on the same line, and without loss of generality, that the police is on your right; then the police moves up, down and left with probability 1/3. Similarly, when the police and you are on the same column, and when your are above the police, then the police moves up, right and left with probability 1/3. When the police and you are not on the same line, nor on the same column, say the police is below you and on your right, then the police moves up and left with probability 1/2. Your objective is to maximise your average discounted reward.

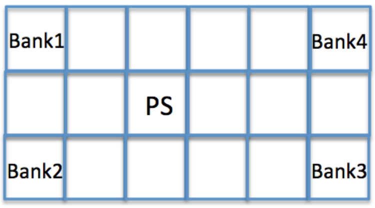

Figure 2: The city

1. Formulate the problem as an MDP.
2. Solve the problem, and display the value function and an optimal policy as a function of λ.