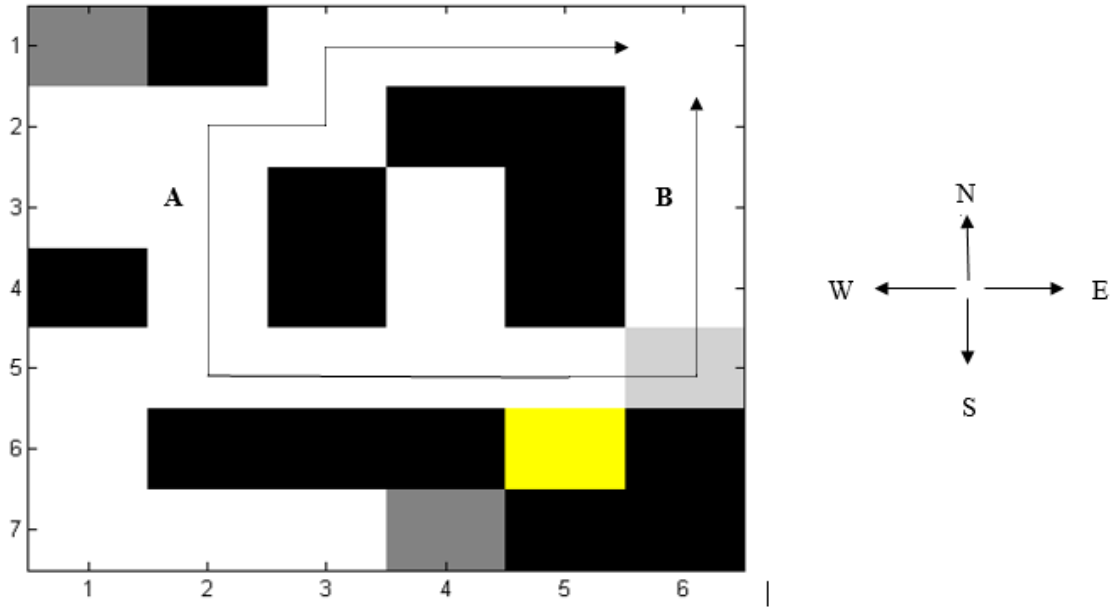
I constructed a two-dimensional maze for the agent to navigate in (see figure). The starting position is at coordinate (5, 6) marked in yellow. The purpose of the agent is to exit (6, 1) the maze via a route that yields an optimal payoff. By optimal here I mean the minimum punishment/maximum reward possible.



The agent has a choice of routes, route A and route B. The former leads to the exit via a long detour whilst the latter leads to the exit via a much shorter (but riskier) route. There is a trap along route B which forces the agent to make decision based on its prior knowledge of the incidence of the traps. The occurrence of the trap within the maze was modelled through a Poisson process. By varying the parameters of the Poisson process I was able to set the risk involved in taking this specific route which led to the exit.