Single Agent Planning

Lab Assignment 1
Autonomous Agents
Master Artificial Intelligence
University of Amsterdam

Camiel Verschoor
StudentID: 10017321
UvAnetID: 6229298
Verschoor@uva.nl

Steven Laan
StudentID: 6036031
UvAnetID: 6036031
S.Laan@uva.nl

Auke Wiggers
StudentID: 6036163
UvAnetID: 6036163
Auke.Wiggers@student.uva.nl

September 19, 2012

1   Introduction

In this report we will discuss the predator versus prey world Markov Decision Process (MDP). The basic environment of the predator versus prey MDP is a 11 by 11 toroidal grid, which means that the north and south side and the east and west side are connected. The predator and prey can be anywhere on the grid, however they cannot have the same position. In this report the starting position of the prey is $(5, 5)$ and of the predator is $(0, 0)$.

In the MDP the prey behaves in a known way, and therefore can be part of the environment. The prey has five actions, namely, $ACTION\_NORTH$, $ACTION\_EAST$, $ACTION\_WEST$, $ACTION\_WEST$ and $ACTION\_STAY$. The prey performs action $ACTION\_STAY$ with probability $P(ACTION\_STAY) = 0.8$ and performs randomly any of the other actions with equal probability.

The predator is the agent in the MDP and has the same actions as the prey. The goal of the predator is to catch the prey, which occurs when the predator stands next to the prey and performs an action towards the prey. The reward for catching the prey is 10 and the immediate reward for any other action is 0.

This report will focus on the planning scenario, which means that the entire MDP is known to the agent. Therefore, the agent can determine the optimal policy even before iterating over the environment.