

# Autonomous Agents Report

Agnes van Belle (10363130),  
Maaïke Fleuren (10350470),  
Norbert Heijne (10357769),  
Lydia Mennes (10333843)

September 20, 2012

This report has been written for the Master Artificial Intelligence course Autonomous Agents. The assignments that will be worked out in detail in this report include the following topics: "Single Agent Planning, Single Agent Learning, and Multi-Agent Planning and Learning". In this report motivation will be given for the design choices and specific programming choices that have been made. The explanation and motivation follows the must-have (depicted by **M**) and should/could-have (depicted by **SC**) structure that has been set up for the assignments. By doing so we hope to make our report easy to read for the teachers/assistants that will be grading this.

## 1 Assignment 1: Single Agent Planning

### 1.1 (M) Simulating the Environment

The choice has been made to not encode the positions of the agents as part of a grid, i.e. matrix. Instead the agents both know their own position and on each iteration the position of the other agent is given as input by the environment, as we have stated in the Agent interface. This is needed for prey to see if the predator is next to it, to prevent it moving towards the prey. In the predator's case it might be necessary later on to know the position of the prey although it is not necessary for this particular sub-assignment.

As part of the assignment a mean and a standard deviation was asked for 100 runs with the use of the random policy for the predator's behaviour. The lowest amount of time steps observed was 19 time steps. The optimal amount of time steps given that the prey would remain still throughout the trial run would be 10. The highest amount observed was 1194 steps. The average amount of time steps was 296.93 time steps and the standard deviation was 244.55 time steps (rounded up). This is a clear indication of the inefficiency of the random policy in this particular setting.

### 1.2 (SC) Iterative Policy Evaluation

### 1.3 (SC) State space reduction

### 1.4 (M) Value Iteration

0.000000	0.000002	0.000011	0.000074	0.000438	0.001730	0.000438	0.000074	0.000438	0.000074	0.000011	0.000002	0.000000
0.000002	0.000011	0.000075	0.000498	0.003195	0.013773	0.003195	0.000498	0.003195	0.000498	0.000075	0.000011	0.000002
0.000011	0.000075	0.000498	0.003443	0.021739	0.117564	0.021739	0.003443	0.021739	0.003443	0.000498	0.000075	0.000011
0.000074	0.000498	0.003443	0.021739	0.166976	0.816327	0.166976	0.021739	0.816327	0.021739	0.003443	0.000498	0.000074
0.000438	0.003195	0.021739	0.166976	0.816327	10.000000	0.816327	0.166976	10.000000	0.816327	0.003195	0.000498	0.000438
0.001730	0.013773	0.117564	0.816327	10.000000	0.000000	0.000000	0.816327	0.000000	0.117564	0.013773	0.001730	0.000438
0.000438	0.003195	0.021739	0.166976	0.816327	10.000000	0.816327	0.166976	0.816327	0.021739	0.003195	0.000438	0.000074
0.000074	0.000498	0.003443	0.021739	0.166976	0.816327	0.166976	0.021739	0.816327	0.003443	0.000498	0.000074	0.000011
0.000011	0.000075	0.000498	0.003443	0.021739	0.117564	0.021739	0.003443	0.117564	0.000498	0.000075	0.000011	0.000002
0.000002	0.000011	0.000075	0.000498	0.003195	0.013773	0.003195	0.000498	0.013773	0.000498	0.000011	0.000002	0.000000
0.000000	0.000002	0.000011	0.000074	0.000438	0.001730	0.000438	0.000074	0.000438	0.000074	0.000011	0.000002	0.000000

## 1.5 (SC) Policy Iteration

## 2 Appendix

Assignment 1 first **M**, 100 runs for the random policy predator.

Timesteps:421  
Timesteps:831  
Timesteps:476  
Timesteps:74  
Timesteps:537  
Timesteps:40  
Timesteps:468  
Timesteps:465  
Timesteps:105  
Timesteps:123  
Timesteps:227  
Timesteps:658  
Timesteps:696  
Timesteps:153  
Timesteps:426  
Timesteps:431  
Timesteps:24  
Timesteps:197  
Timesteps:517  
Timesteps:313  
Timesteps:492  
Timesteps:213  
Timesteps:457  
Timesteps:392  
Timesteps:47  
Timesteps:178  
Timesteps:459  
Timesteps:624  
Timesteps:881  
Timesteps:100  
Timesteps:244  
Timesteps:127  
Timesteps:213  
Timesteps:145  
Timesteps:45  
Timesteps:301  
Timesteps:628  
Timesteps:248  
Timesteps:88  
Timesteps:123  
Timesteps:82  
Timesteps:206  
Timesteps:181  
Timesteps:771  
Timesteps:114  
Timesteps:238  
Timesteps:118  
Timesteps:67  
Timesteps:41  
Timesteps:662

Timesteps:27  
Timesteps:73  
Timesteps:217  
Timesteps:269  
Timesteps:382  
Timesteps:60  
Timesteps:205  
Timesteps:64  
Timesteps:133  
Timesteps:232  
Timesteps:148  
Timesteps:504  
Timesteps:113  
Timesteps:316  
Timesteps:151  
Timesteps:178  
Timesteps:53  
Timesteps:526  
Timesteps:150  
Timesteps:690  
Timesteps:490  
Timesteps:116  
Timesteps:288  
Timesteps:79  
Timesteps:163  
Timesteps:266  
Timesteps:566  
Timesteps:1194  
Timesteps:133  
Timesteps:690  
Timesteps:136  
Timesteps:121  
Timesteps:123  
Timesteps:492  
Timesteps:288  
Timesteps:185  
Timesteps:19  
Timesteps:78  
Timesteps:250  
Timesteps:42  
Timesteps:268  
Timesteps:190  
Timesteps:231  
Timesteps:393  
Timesteps:338  
Timesteps:100  
Timesteps:49  
Timesteps:653  
Timesteps:1173  
Timesteps:421  
Average timesteps over 100 trials: 296.93  
Standard deviation over 100 trials: 244.54689754728025