

CSE4082 – Assignment #1

(Due 17.11.2023 at 23:59, electronic submission only (report and source code, no executable!))

Consider an extended vacuum cleaner world in which there are three locations, namely rooms A, B, and C. The geography of the environment is as follows:

Room A	Room B	Room C
--------	--------	--------

Each room may have two different states that are *clean* and *dirty*. Initially, all three rooms are dirty, and the robot is located at Room B. The vacuum agent perceives which room it is in and whether the room is dirty or clean. At any given time step, the agent can choose only one of the following actions: to move left (*left*), move right (*right*), suck up the dirt (*suck*), or do nothing (*no-op*).

a) Assume that if at any given time step: either

(1) room X is clean; or

(2) room X is dirty, the agent is in room X and it performs *suck* action;

then at the beginning of the next time step, with probability P_x , room X gets dirty if it is clean (X is one of A, B, or C).

i) Propose a utility-based agent. Assume the agent is rewarded one point for each clean room at the end of each time step. Call this agent Agent A.

ii) Assume that the agent is rewarded one point for each clean room at the end of each time step and penalized 0.5 point for each move action. Modify the proposed Agent A accordingly and call this agent Agent B.

iii) Simulate the performances of the proposed Agent A and Agent B for 1000 time steps.

Your program should get three inputs from command-line: utility values P_a , P_b and P_c ; and provide two output text files in the following format:

a.txt: State – Action information for Agent A. The format of this file would be

B, D, D, D (The robot is in Room B and all three rooms are dirty.)

suck

1 (Current score)

B, D, C, D

right

C, D, C, D

2

...

Using this format, the last line of this file will show the total score of the corresponding agent.

b.txt: State – Action information for Agent B. The format of this file would be the same as a.txt

iv) For each configuration below, you should run the simulation 10 times and report the mean and the standard deviation of the rewarded points of the agents.

Configuration	P_A	P_B	P_C
1	0.3	0.3	0.3
2	0.5	0.2	0.1
3	0.2	0.4	0.2
4	0.5	0.1	0.3
5	0.5	0.3	0.8

Do not forget to seed the random generator using the time function. Note that the agents should **NOT** know the actual probability values P_A , P_B , and P_C . They will try to learn.

v) Explain the algorithms of the agents and briefly comment on the results of the simulation.

Notes:

- This assignment can be done in groups of two or three.
- Any programming language can be used.
- You should provide a design document and output text files for the above configurations. The design document should include the algorithms of the agents and discussion of the simulation results.