ECE8843 Assignment 2 : Learning Robot Control Strategies

Ana Huamán Quispe

I. Problem Statement

Robby, the Soda-Can-Collecting robot has the job of cleaning up his environment by collecting soda cans. His gridded world is shown in Fig.I:
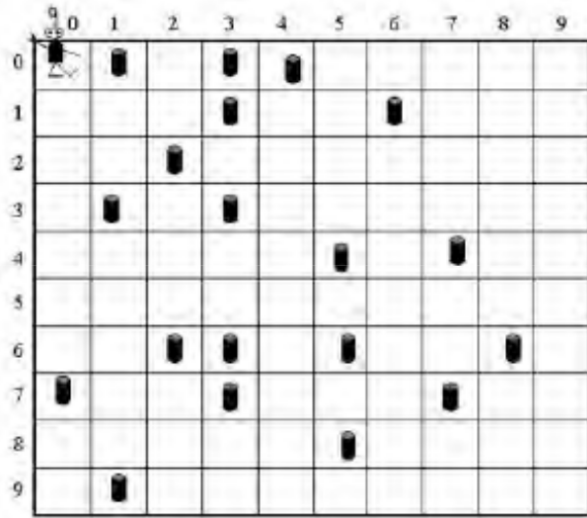

Fig. 1. 10×10 gridded world

A. Actions

The robot has 7 possible actions:

- *North:* Move one grid up
- *South:* Move one grid down
- *East:* Move one grid right
- *West:* Move one grid left
- *Stay:* Stay on same location
- *Pick up:* Stay, bend and pick up can (if any)
- *Random:* Any of the above

The number of allowed actions per episode is 200.

B. Rewards

- *Move (North, South, East or West):* -1
- *Bump wall:* -5
- *Failed pick up:* -2
- *Successful pick up:* +10

The author is with the Georgia Institute of Technology, Atlanta, GA.ahuaman3@gatech.edu

C. States

Robot states are composed of 5 variables, each of which can have 3 possible values:

- *Current Grid:* Free, Can or Wall
- *North Grid:* Free, Can or Wall
- *South Grid:* Free, Can or Wall
- *East Grid:* Free, Can or Wall
- *West Grid:* Free, Can or Wall

Having a total of $3^5 = 243$ states

II. Greedy Approach

We solved the problem by using a simple variant of the well-known $\epsilon - greedy$ approach. At each time step, the robot has two possible ways to choose an action:

- *With probability ϵ:* Choose a random action such that the robot does not bump into walls (north, south, east, west, stay or pick up)
- *With probability $1 - \epsilon$:* Choose the action with highest reward, that is:
  - If *CURRENT_GRID* is *CAN* $\rightarrow$ Action is *PICK_UP*
  - Else
    * If *NORTH_GRID* is *CAN* $\rightarrow$ Action is *NORTH*
    * If *SOUTH_GRID* is *CAN* $\rightarrow$ Action is *SOUTH*
    * If *EAST_GRID* is *CAN* $\rightarrow$ Action is *EAST*
    * If *WEST_GRID* is *CAN* $\rightarrow$ Action is *WEST*
    * Else $\rightarrow$ Action is *RANDOM*

Although fairly simple, this approach works moderately well. Table I shows the results obtained in 10 runs of our program with randomly generated starting positions:

TABLE I
RESULTS OF GREEDY APPROACH ($\epsilon = 0.05$)

Episode	Start Location	Final Value	Cans Picked
0	$(3, 1)$	-38	15
1	$(0, 2)$	-39	15
2	$(8, 3)$	6	19
3	$(6, 4)$	-35	15
4	$(0, 3)$	-58	13
5	$(8, 2)$	-62	13
6	$(8, 8)$	-17	17
7	$(8, 5)$	6	19
8	$(1, 3)$	-69	12
9	$(8, 4)$	-16	17

III. Sample Trial Run

To exemplify the usefulness of our approach we provide the result of 01 run of our code. The output is attached on the zip file containing this .pdf (name of the document is sampleRun.txt)