The line follower robot Using Reinforcement Q-Learning Algorithm

Team members:

1- Solwan Shokry Ahmed Mohamed.

2- Rana Mohamed Hussein Mohamed.

3- Sarah Gamal El-Deen Mohamed.

4- Rahma Mohamed Makram Saeid.

5- Amira Abdelfattah Abdelrazik.

Date

06/07/2021

Course title

Machine Learning – Microcontroller systems

Supervised by: Dr/ Amr El-Sayed

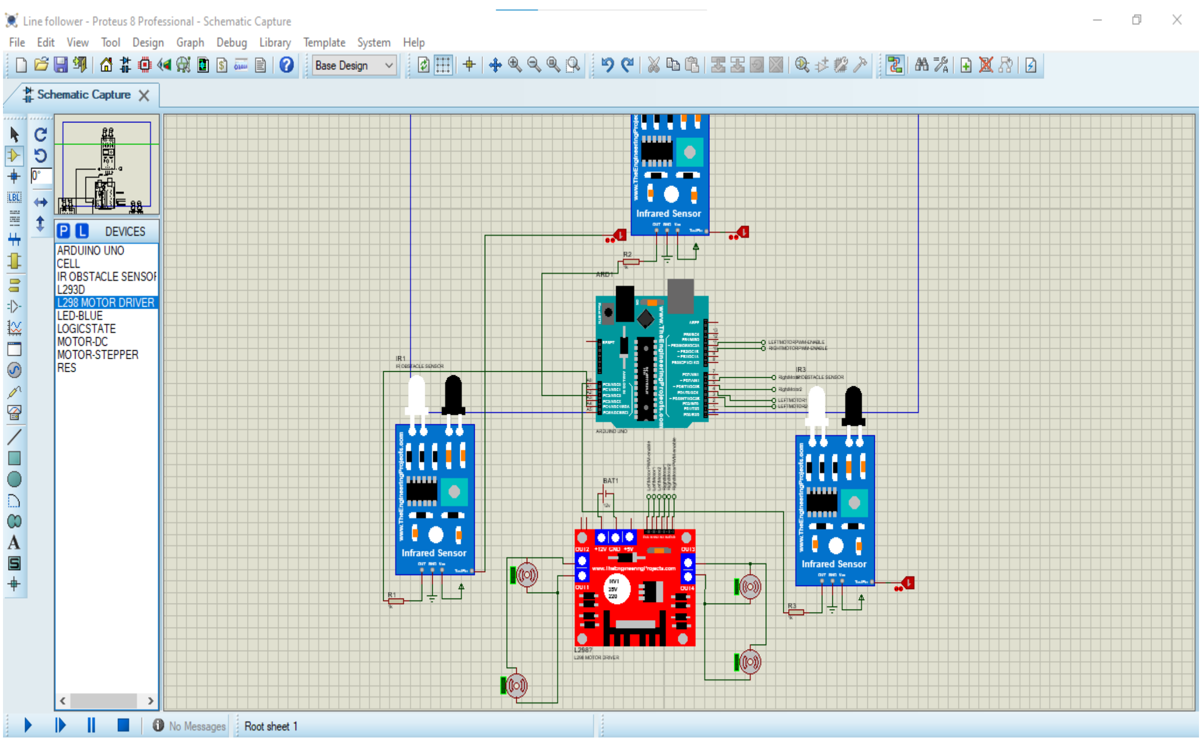
Line Follower Robot

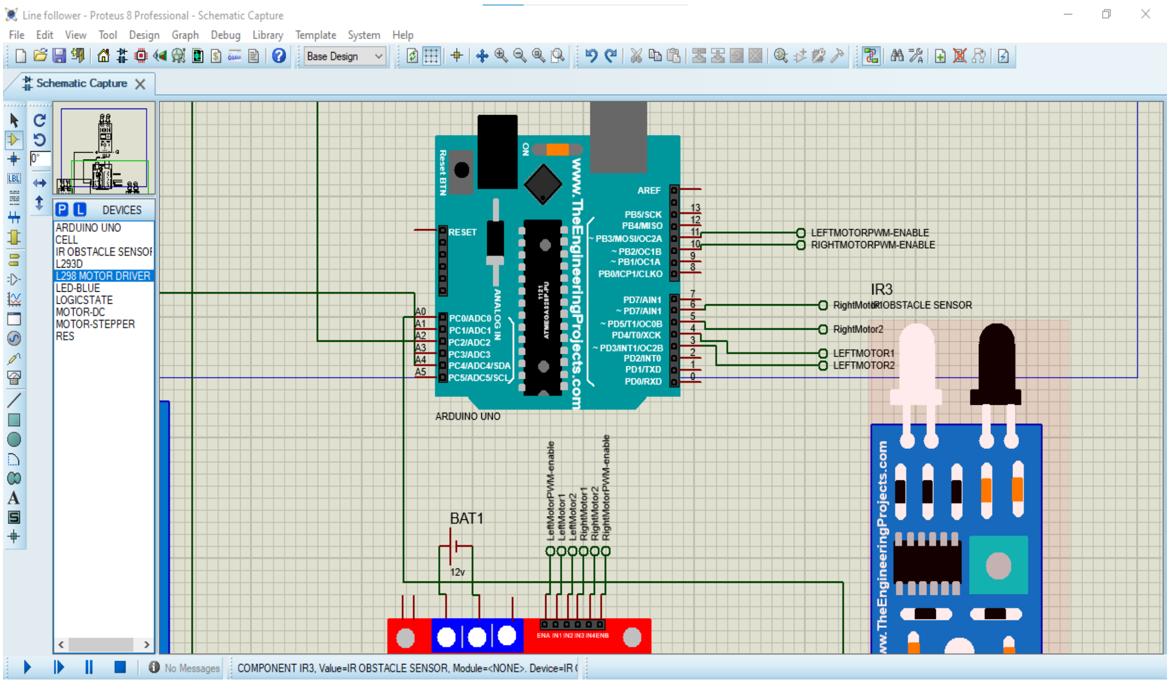
1- Microcontroller & Embedded System Features

We used ATmega328p Microcontroller, embedded on ARDUINO UNO kit. The circuit's operation requires three Infrared sensors to detect the black line's location. The sensor's analog readings are connected to distinct ADC channel. The ATmega328p's Analog to Digital Converter then converts them into Digital readings used in the Robot's functionality Algorithm described below.

The robot is derived by four DC-motor wheels. Which are connected to the ARDUINO UNO kit through L298N Motor Driver too control the motor (wheels) motion Direction. It's controlled through the controller's internal 8-bits Timer0A by supplying it with Pulse wave to control the wheels (motor) speed.

Schematic Diagram:

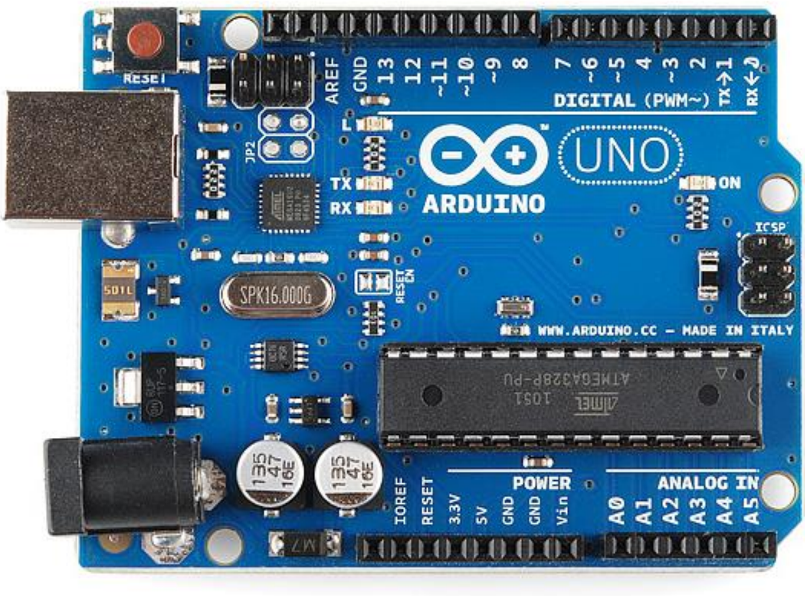




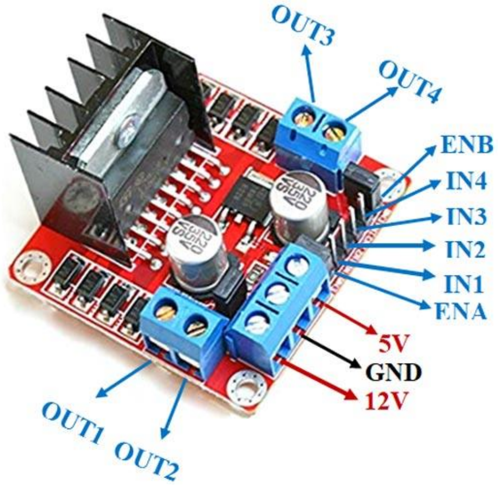
- Hardware Components:

1. ARDUINO UNO kit:

Having Power (USB / Barrel Jack), Pins (5V, 3.3V, GND, Analog, Digital, PWM, AREF), Reset Button, Power LED Indicator, TX RX LEDs, ATmega328p MCU IC, Voltage Regulator.



2. L298N Motor Driver Module:



This **L298N Motor Driver Module** is a high power motor driver module for driving DC and Stepper Motors. This module consists of an L298 motor driver IC and a 78M05 5V regulator. **L298N Module** can control up to 4 DC motors, or 2 DC motors with directional and speed control.

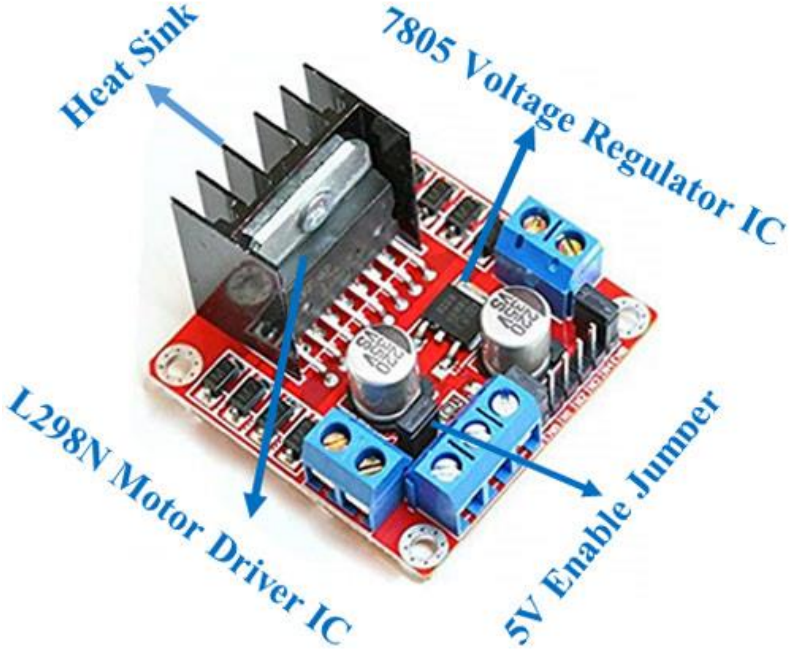
L298N Module Pin Configuration:

Pin Name	Description
IN1 & IN2	Motor A input pins. Used to control the spinning direction of Motor A
IN3 & IN4	Motor B input pins. Used to control the spinning direction of Motor B
ENA	Enables PWM signal for Motor A
ENB	Enables PWM signal for Motor B
OUT1 & OUT2	Output pins of Motor A
OUT3 & OUT4	Output pins of Motor B
12V	12V input from DC power Source
5V	Supplies power for the switching logic circuitry inside L298N IC
GND	Ground pin

L298 Module Features & Specifications:

- Driver Model: L298N 2A
- Driver Chip: Double H Bridge L298N
- Motor Supply Voltage (Maximum): 46V
- Motor Supply Current (Maximum): 2A
- Logic Voltage: 5V
- Driver Voltage: 5-35V
- Driver Current:2A
- Logical Current:0-36mA
- Maximum Power (W): 25W
- Current Sense for each motor
- Heatsink for better performance
- Power-On LED indicator

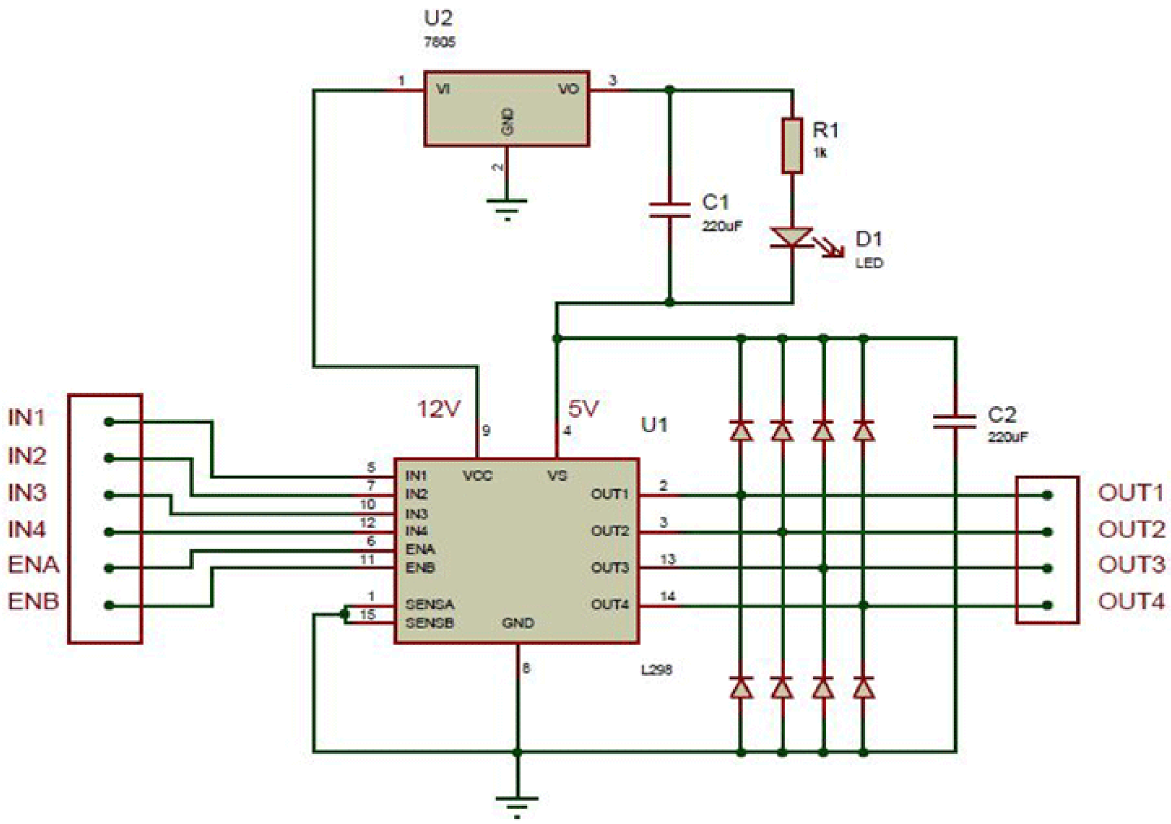
The L298N Motor Driver module consists of an L298 Motor Driver IC, 78M05 Voltage Regulator, resistors, capacitor, Power LED, 5V jumper in an integrated circuit.



78M05 Voltage regulator will be enabled only when the jumper is placed. When the power supply is less than or equal to 12V, then the internal circuitry will be powered by the voltage regulator and the 5V pin can be used as an output pin to power the microcontroller. The jumper should not be placed when the power supply is greater than 12V and separate 5V should be given through 5V terminal to power the internal circuitry.

ENA & ENB pins are speed control pins for Motor A and Motor B while IN1& IN2 and IN3 & IN4 are direction control pins for Motor A and Motor B.

Internal circuit diagram of L298N Motor Driver module is given below:



Applications of L298N Module

- Drive DC motors.
- Drive stepping motors
- In Robotics

3. <u>IR Sensor:</u>

IR Sensor Working Principle:

There are different types of infrared transmitters depending on their wavelengths, output power and response time. An IR sensor consists of an IR LED and an IR Photodiode, together they are called as PhotoCoupler or OptoCoupler.

IR Transmitter or IR LED:

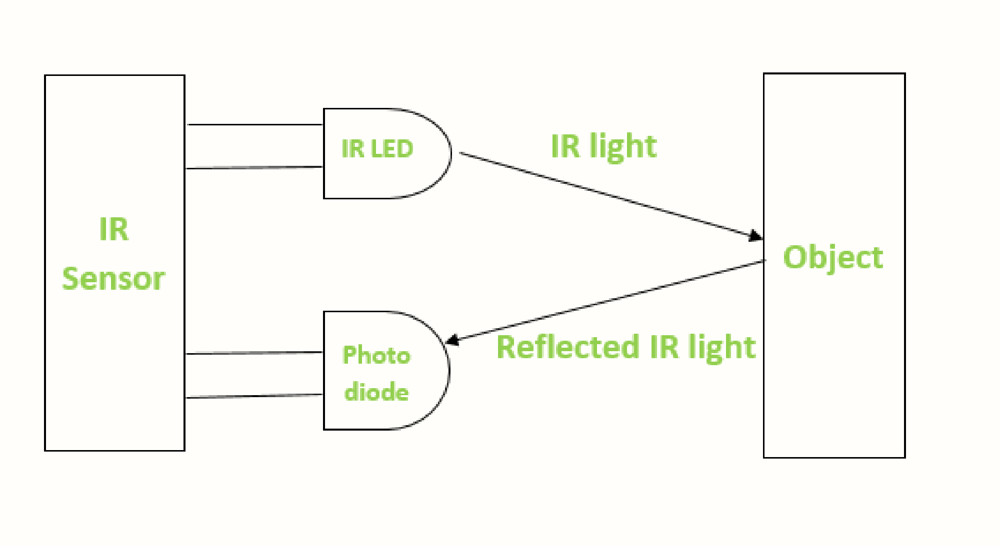
Infrared Transmitter is a light emitting diode (LED) which emits infrared radiations called as IR LED's. Even though an IR LED looks like a normal LED, the radiation emitted by it is invisible to the human eye.
The picture of an Infrared LED is shown below.

IR Receiver or Photodiode:

Infrared receivers or infrared sensors detect the radiation from an IR transmitter. IR receivers come in the form of photodiodes and phototransistors. Infrared Photodiodes are different from normal photo diodes as they detect only infrared radiation. Below image shows the picture of an IR receiver or a photodiode,

Different types of IR receivers exist based on the wavelength, voltage, package, etc. When used in an infrared transmitter – receiver combination, the wavelength of the receiver should match with that of the transmitter.

The emitter is an IR LED and the detector is an IR photodiode. The IR photodiode is sensitive to the IR light emitted by an IR LED. The photo-diode's resistance and output voltage change in proportion to the IR light received. This is the underlying working principle of the IR sensor.



When the IR transmitter emits radiation, it reaches the object and some of the radiation reflects back to the IR receiver. Based on the intensity of the reception by the IR receiver, the output of the sensor defines.

4. Dc-motor Wheels:

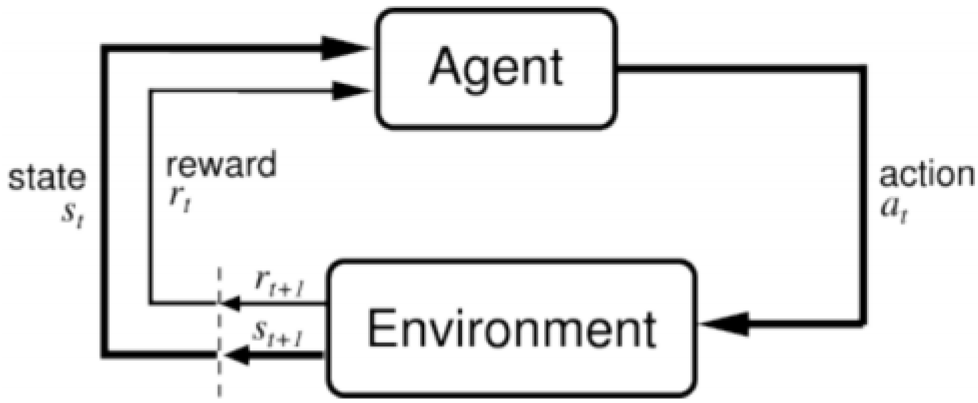
Female jumpers
Soldering Iron
Soldering tin
Bread board
Wooden Base

2- Machine Learning Part

- Reinforcement Q-Learning Algorithm model:

The key entities of interest are the environment, action, reward and state. Let's give ourselves a name as well - we are agents in this whole game of life. This whole paradigm of exploring our lives and learning from it through actions, rewards and states establishes the foundation of reinforcement learning. In fact, this is almost how we act in any given circumstance in our lives

For a robot, an environment is a place where it has been put to be used. This robot is itself the agent. For example, The tasks we discussed just now, have a property in common - these tasks involve an environment and expect the agents to learn from that environment. This is where traditional machine learning fails and hence the need for reinforcement learning.



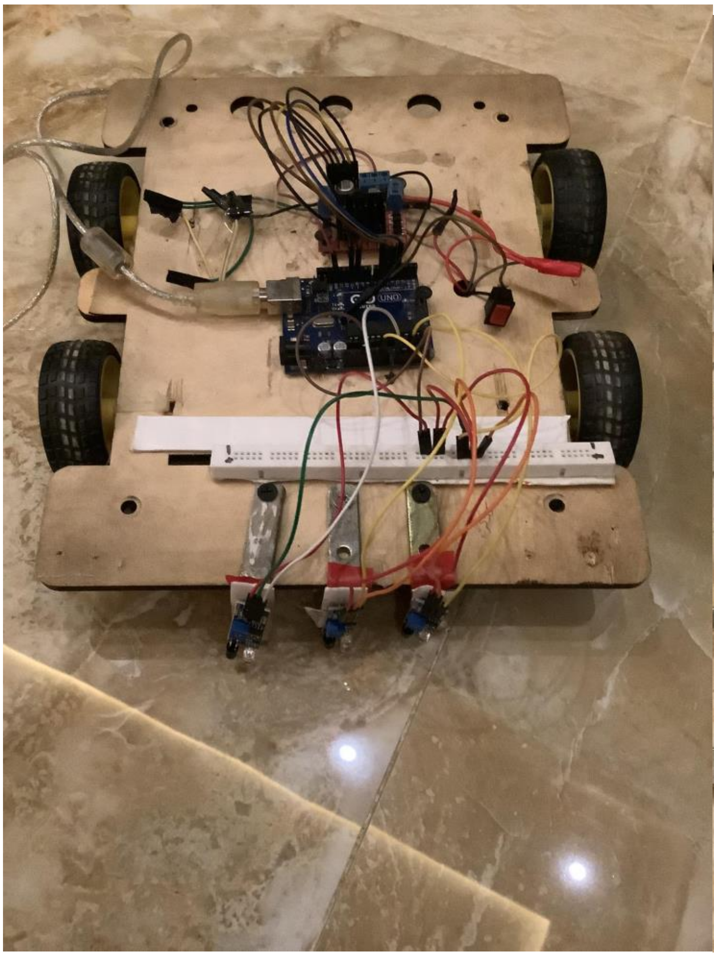
Defining the problem statement:

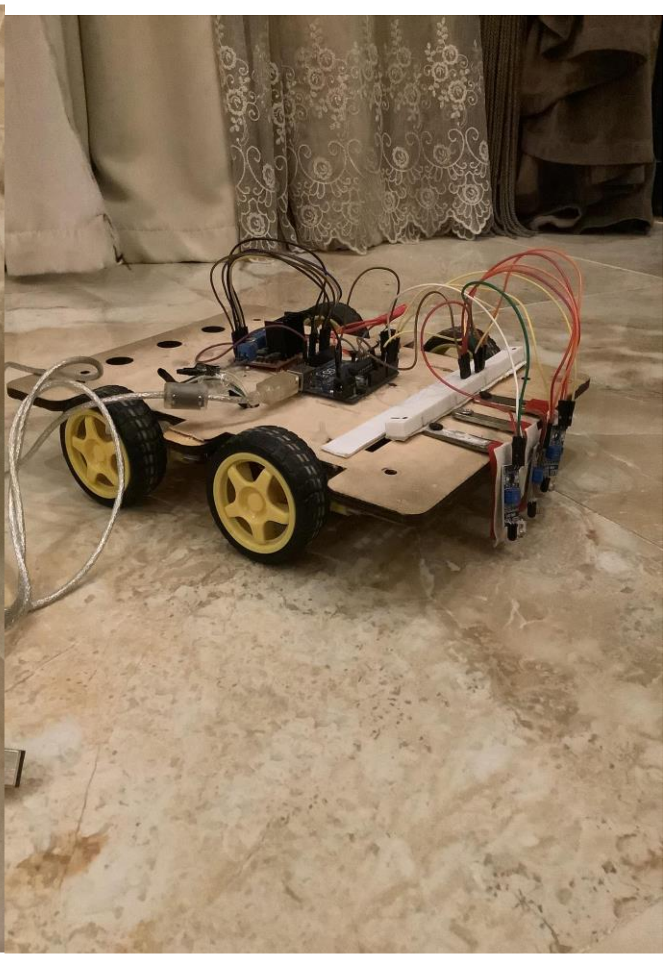
1- Environment:

A white surface with black line track drawn over it for the robot to follow

2- Agent:

Our line follower Robot





3- States:

In our case the states depends on the readings of the IR sensors placed on Right, Center, and Left side to side and fixed to the front of the robot. The sensors readings are what determines the states that the robot has to decide on what is the next Action to take where:

The infrared sensor will output 1 when it detects the black line, having three sensors gives 2^3 = 8 possible cases:

	Left	Center	Right
0:	0	0	0
1:	0	0	1
2:	0	1	0
3:	0	1	1
4:	1	0	0
5:	1	0	1
6:	1	1	0
7:	1	1	1

Stop: there is no right nor ledt nor forward black line to follow

State 1: Go right is the only valid option

State2: Move Forward is the only valid option

State3: Move forward & move Right are valid

State4: Move Left is the only valid option

State5: Move Left or move Right are the valid options

State6: Move Left or Move forward

State7: Forward & Left & Right are all valid options

4- Actions:

The robot’s actions are one of four:

- Move Forward
- Turn Right
- Turn Left
- Stop

It is to determine the action to take depending on the Q value calculated by the

$$\boxed{Q(s,a) \leftarrow (1-\alpha)\,Q(s,a) + \alpha\,[r + \gamma \max_{a'} Q(s',a')]}$$

In our Case we initialized the Q[States_count][Actions_count] with all zero’s as the robot has no pre-stored training values. It starts learning as soon as it powers up then it updated the value of Q for each state and its corresponding action using this formula.

5- Rewards:

We created a 2D Array of size 8*4 Rewards[States_count][Actions_count]
to store the values of Rewards calculated for each (S,a) pair, initialized with all 0's
and setting the LEARNING RATE (Alpha) to 0.1
and the Discount Factor (Gama) to 0.5

- **Q-Learning Algorithm:**

Initialize $Q(s,a)$ arbitrarily

Repeat (for each episode):

  Initialize s

  Repeat (for each step of episode):

    Choose a from s using policy derived from Q

    Take action a, observe r, s'

    Update

$$Q(s,a) \leftarrow Q(s,a) + \alpha[r + \gamma \max_{a'} Q(s',a') - Q(s,a)]$$

    $s \leftarrow s'$;

  Until s is terminal