All programs of Problem 3 Part 1 & 2 are implemented with policy and value grid. Here are the output of all at the 200th iteration:-

[Value iterations contain expectation value of reward and policy iteration board contains best policy move (U,D,L,R,X for up,down,left,right and no move respectively, empty block is the wall in policy iteration board]

BOARD STRATEGIES FOR DIFFERENT MAGNETO WITH DIFFERENT POLICIES

Active Magneto [value iteration]

Final Values:
```
----------------------------------
 10.44| 12.28| 14.45| 20.00| 20.00|
----------------------------------
 10.44| 14.45| 12.28| 14.45|-17.00|
----------------------------------
 14.45| 17.00| 12.28| 12.28| 3.94|
----------------------------------
 17.00| 20.00| 0.00|-20.00| 0.00|
----------------------------------
 0.00| 17.00| 20.00| 14.45| 0.00|
```

Final Policy (D,U,L,R,X for down,up,left,right and no move:)
```
-------------------------------
 R | R | L | R | X |
-------------------------------
 D | D | U | U | U |
-------------------------------
 D | D | L | U | X |
-------------------------------
 R | D |   | U | U |
-------------------------------
 U | X | L | L | L |
```

Active Magneto [Policy iteration]

Values 200:
```
------------------------------------
 10.44| 12.28| 10.44|-20.00| 17.00|
------------------------------------
 12.28| 10.44| 12.28| 10.44| 20.00|
------------------------------------
 12.28| 12.28| 14.45| 12.28|-0.13|
------------------------------------
 14.45| 14.45| 0.00| 10.44|-0.13|
------------------------------------
 14.45| 20.00| 0.00| 7.54| 6.41|
```

Policy 200:
```
-----------------------------
 D | L | D | R | D |
-----------------------------
 U | U | D | L | X |
-----------------------------
 D | D | X | L | U |
-----------------------------
 D | D |   | U | D |
-----------------------------
 R | X | L | U | L |
```

Lazy Magneto [Policy Iteration]

Values 200:
```
-----------------------------------
-5.00|-0.62| 10.00| 20.00| 10.00|
-----------------------------------
-10.00|-0.31|-0.62| 10.00|-0.62|
-----------------------------------
-10.00|-0.62|-20.00|-10.00|-1.25|
-----------------------------------
-10.00| 10.00| 0.00|-0.00|-0.00|
-----------------------------------
 2.50|-10.00|-20.00|-0.00|-0.00|
```

Policy 200:

```
------------------------------
 D | R | R | X | L |
------------------------------
 U | X | U | U | L |
------------------------------
 R | D | U | U | U |
------------------------------
 R | D |   | D | D |
------------------------------
 R | X | L | R | U |
------------------------------
 X | X | L | L | L |
```

Lazy Magneto [Value Iteration]

Final Values:

```
-----------------------------------
 12.28| 14.45| 17.00| 17.00| 20.00|
-----------------------------------
 10.44| 12.28| 14.45| 17.00| 17.00|
-----------------------------------
 10.44| 10.44| 12.28| 14.45| 17.00|
-----------------------------------
 17.00| 14.45| 0.00| 14.45| 14.45|
-----------------------------------
 20.00| 17.00| 17.00| 17.00| 14.45|
```

Final Policy (D,U,L,R,X for down,up,left,right and no move:)

```
------------------------------
 D | R | R | R | X |
------------------------------
 U | U | U | U | U |
------------------------------
 D | U | U | U | U |
------------------------------
 D | D |   | D | U |
------------------------------
 X | L | L | L | L |
```