



Online Robot Learning by Reward and Punishment for a Mobile Robot

D. Suwimonteerabuth
P. Chongstitvatana

Chulalongkorn University, Thailand



Aim

A robot that is more flexible in learning a new task.

Method

- A human trainer influences the robot behavior.
- The robot learns from reward and punishment from a human in real-time.
- The factors that influence the learning process.



Mechanism

- A Finite-State Machine is used as robot controller.
- Genetic Algorithms is used to evolve FSM.

Method of experiment

- A simulator is used to find out the appropriate parameters for GA.
- The physical robot is used online in real time to study the learning characteristics.



Tasks

- A rectangular floor 1.5×2.2 m. surrounded by walls.
- The robot is able to detect walls and stays inside the designated floor.
- The floor is painted with two colors: black and white.
- The goal : the robot will learn to stay in one color.

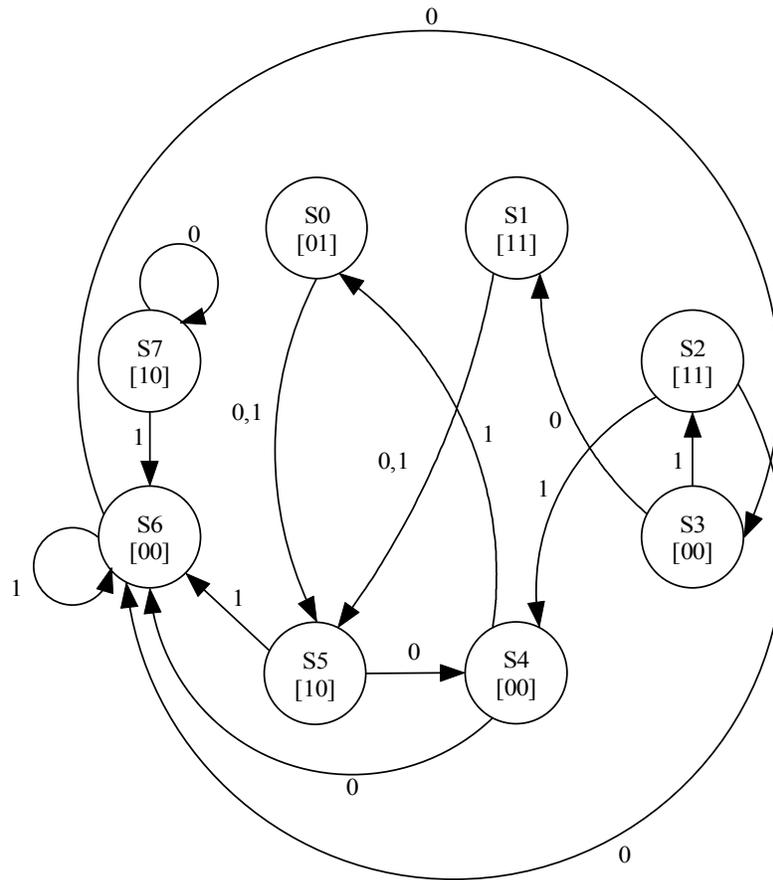


Figure 2: An example FSM encoding

```
as 01 11 11 00 00 10 00 10
    101101 101101 110100
    001010 110000 100110
        011110 111110
```



Problems

Factors that are relevant to the quality of learning.

- effect of reward/punishment
- start position
- maximum number of reward/punish
- size of population
- the human trainer

Real-time experiment

- each FSM runs 30 seconds
- the robot behavior is evaluated by a human trainer
- experiments are performed a couple of times, 60 minutes each



Measurement

The time that it stays in the designated color.

Results

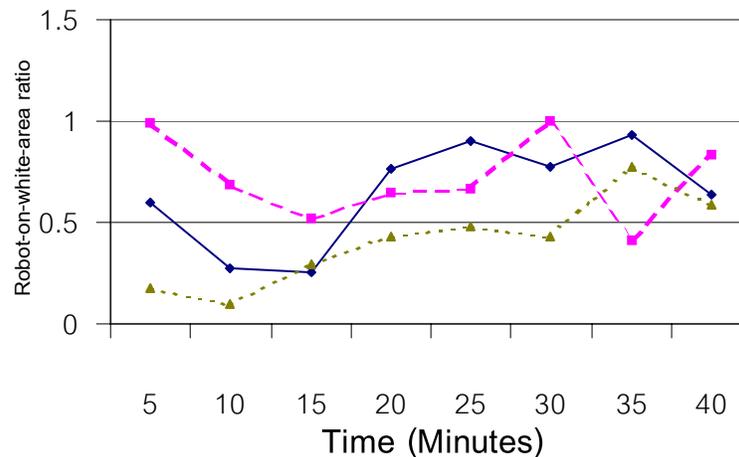


Figure 3: The learning behavior when giving no reward/punishment.

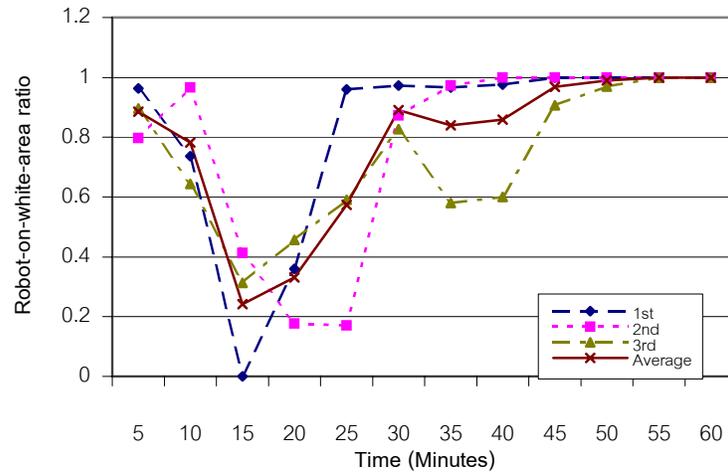


Figure 4: The learning behavior when start position is in white area.

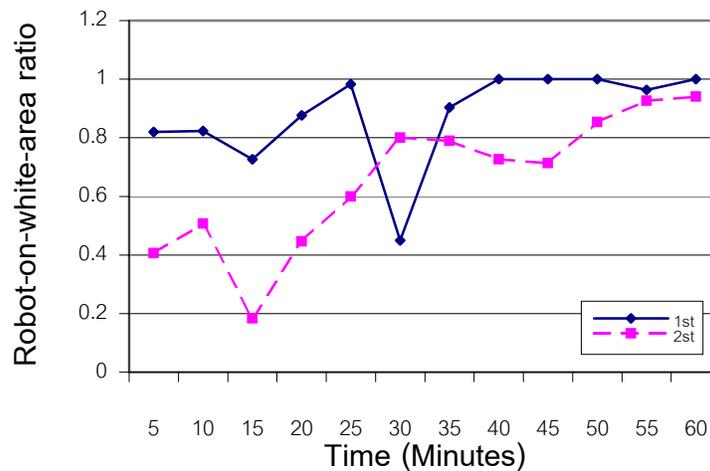


Figure 5: The learning behavior when start position is in black area.

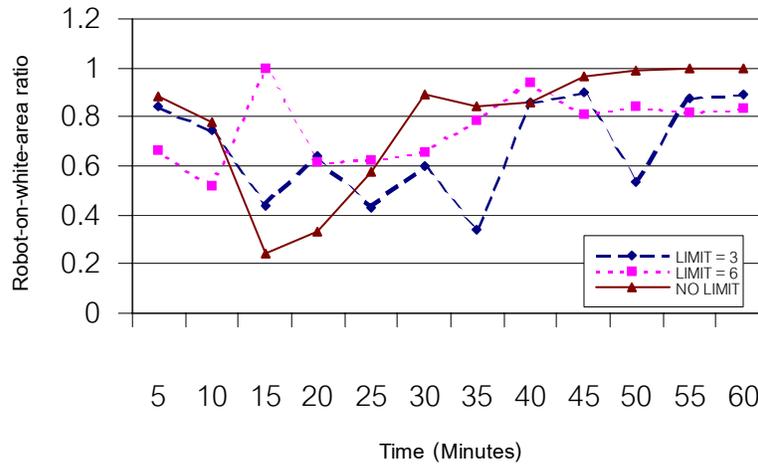


Figure 6: The learning behavior when limiting number of reward/punishment.

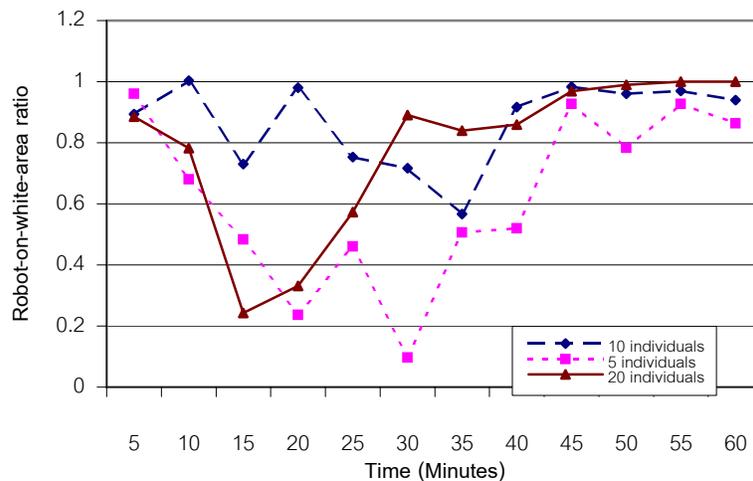


Figure 7: The learning behavior when varying size of population.

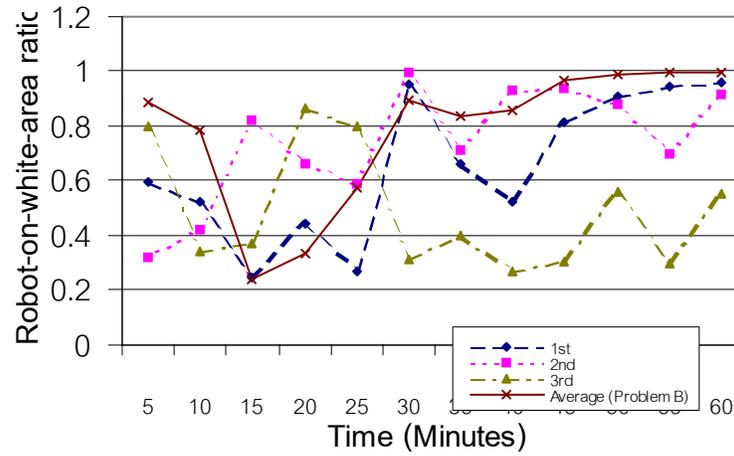


Figure 8: The learning behavior when changing the trainer



Conclusions

- an empirical study of the learning behavior using reward and punishment.
- the robot must learn the behavior that will receive reward and avoid the behavior that will receive punishment.
- Genetic algorithms evolves the controller in the form of a finite state machine.
- Robot is flexible, its behavior can be shaped in real-time and continuously.
- It is still too tedious to train a robot.



Department of Computer Engineering
Chulalongkorn University