

Simulated Smart Car

with Reinforcement Learning

MAROOF H. KHAN

Motivation for using Reinforcement Learning



AlphaGo 2016



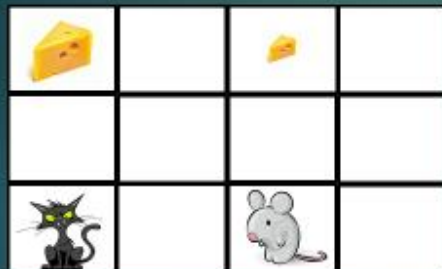
Boston Dynamics 2014



Italian Inst. Tech 2010

- ▶ Start from first principles
- ▶ Update algorithm to increase learning efficiency

What is Reinforcement Learning



100	0	20	0
0	0	0	0
-100	0	Agent	0

- ▶ Movements of the mouse are stochastic
- ▶ Aim is to find a policy [strategy] that maximizes reward of agent [mouse]

Two modes of Learning

Exploration

100		20	
		↑	
-100	←	Agent	→

Exploitation

100		20	← 15
			10 ↑
-100		Agent	→ 5

- ▶ Even if the optimal strategy is found we will have exploration available.
- ▶ After learning is done we have explored all avenues we have optimal strategy [get to cheese in shortest number of steps, avoiding the cat]

Process

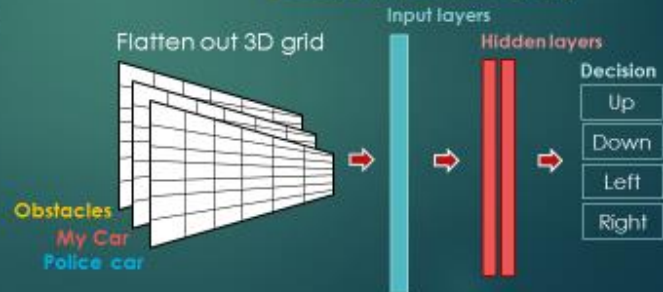


First Implementation

Grid Implementation from scratch



Second Implementation



DEMO

Training



Convention
Training time **15 mins**



Updated
Training time **8 mins**

Static case after training



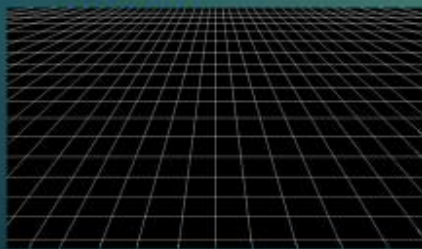
Dynamic Environment Update



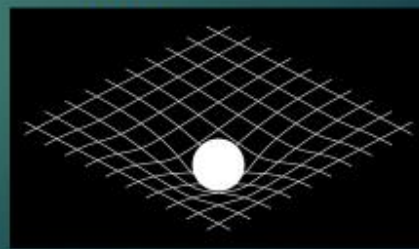
Dynamic Environment Update



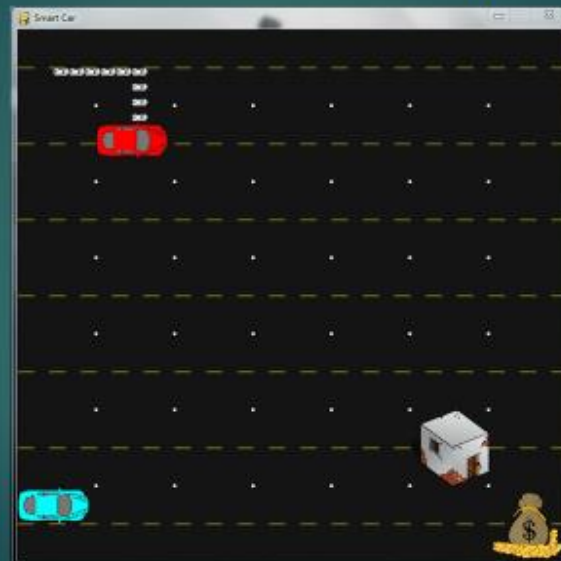
Without Obstacles



With Obstacles



Dynamic Environment Update



Change Reward State without any training



Applications

- ▶ Apply updated algorithm to autonomous robots
- ▶ Use *Inverse RL* to train robot faster

References:

- R. Sutton, A. Barto, Reinforcement Learning: An Introduction
- S. Levine, Z. Popovic, V. Koltun, "Future Construction for Inverse Reinforcement Learning"
- L. Lin, "Programming Robots Using Reinforcement Learning and Teaching", AAAI Proceedings, vol 92, 781-786 (1991)

Maroof H. Khan

Background



PhD, Applied Physics



Technical Analyst
Optical scanner systems

Contact



maroof.hk@gmail.com



[/in/maroofhk](https://www.linkedin.com/in/maroofhk)



github.com/maroofhk



maroofhk.github.io



Technology used:



python

theano



pygame