Intelligent Network Management using Reinforcement Learning

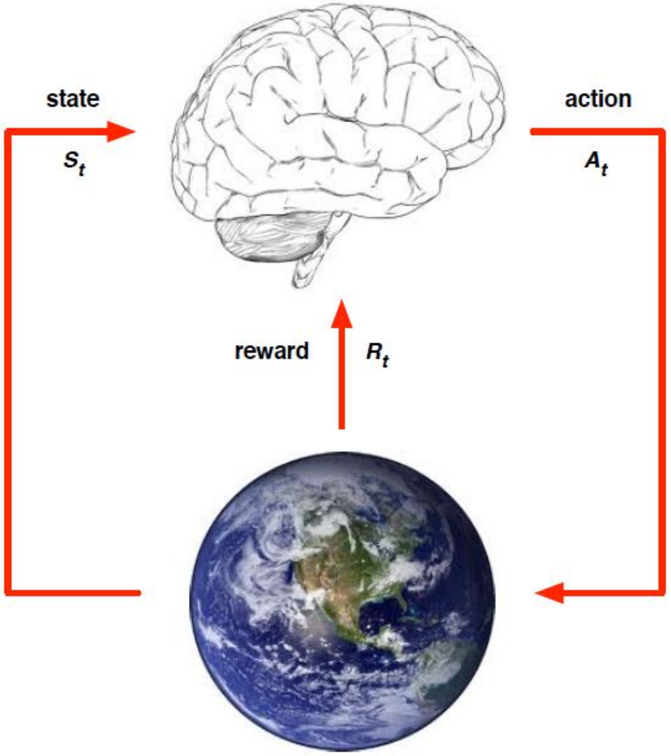
M. Kim, Y. Hong (ETRI)
Y. Han (Koreatech.)
T. Ahn (KT)
K. Kim (ETRI)

NMRG Meeting@IETF 102 – Montreal
2017.07.19

draft-kim-nmrg-rl-03
(https://datatracker.ietf.org/doc/draft-kim-nmrg-rl/)

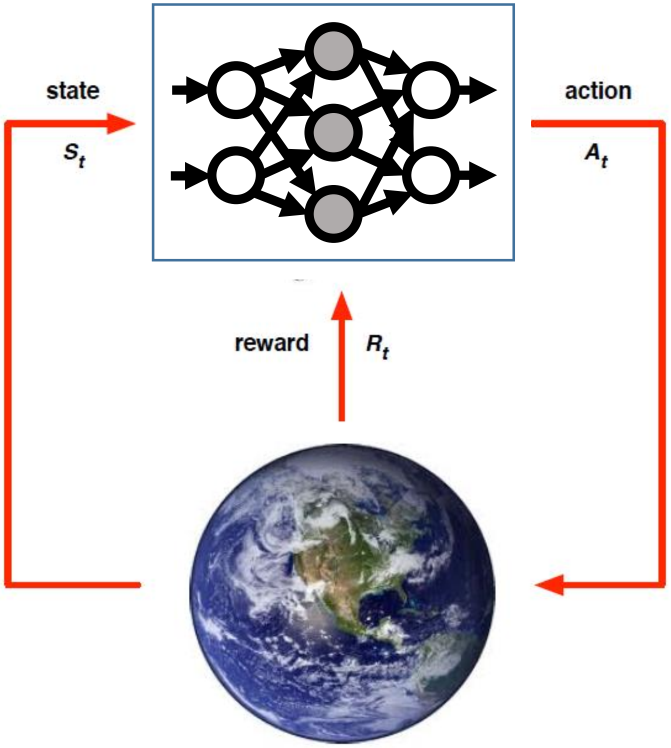
Table of Contents

Reinforcement Learning (RL)



<u>*Tutorial: Deep Reinforcement Learning, David Silver, Google DeepMind*</u>

DRL = RL + Deep Learning

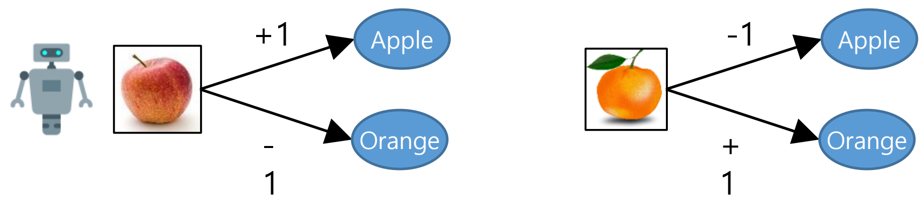


Tutorial: Deep Reinforcement Learning, David Silver, Google DeepMind

Decision Problem

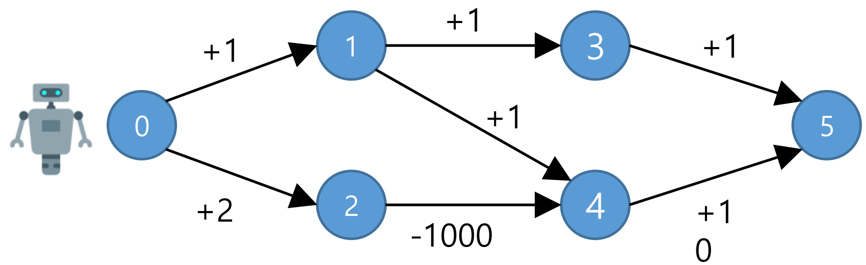
◆ Supervised Learning

- Single Decision



◆ Reinforcement Learning

- Sequential Decision

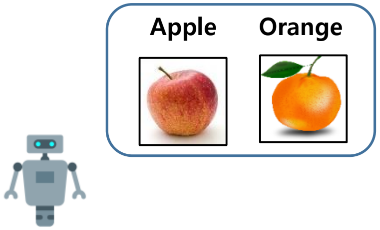


Robot Icon from https://visualpharm.com/free-icons/robot%203-595b40b85ba036ed117da97e
Apple and Orange from https://pixabay.com
Sequential Decision example from 'Reinforcement Learning: A User's Guide, Bill Smart, Washington University in St. Louis.'

Example vs Experience

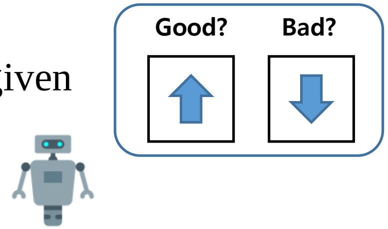
◆ Supervised Learning

- Learning by Example
- Labeled examples are given.
- Problems are defined by ‘examples’



◆ Reinforcement Learning

- Learning by Experience (or Interaction)
- Examples of ‘Good’ and ‘Bad’ (value-based) actions are not given
- Tasks are implicitly given through ‘Reward’



Robot Icon from https://visualpharm.com/free-icons/robot%203-595b40b85ba036ed117da97e
Apple and Orange from https://pixabay.com
Sequential Decision example from ‘Reinforcement Learning: A User’s Guide, Bill Smart, Washington University in St. Louis.’

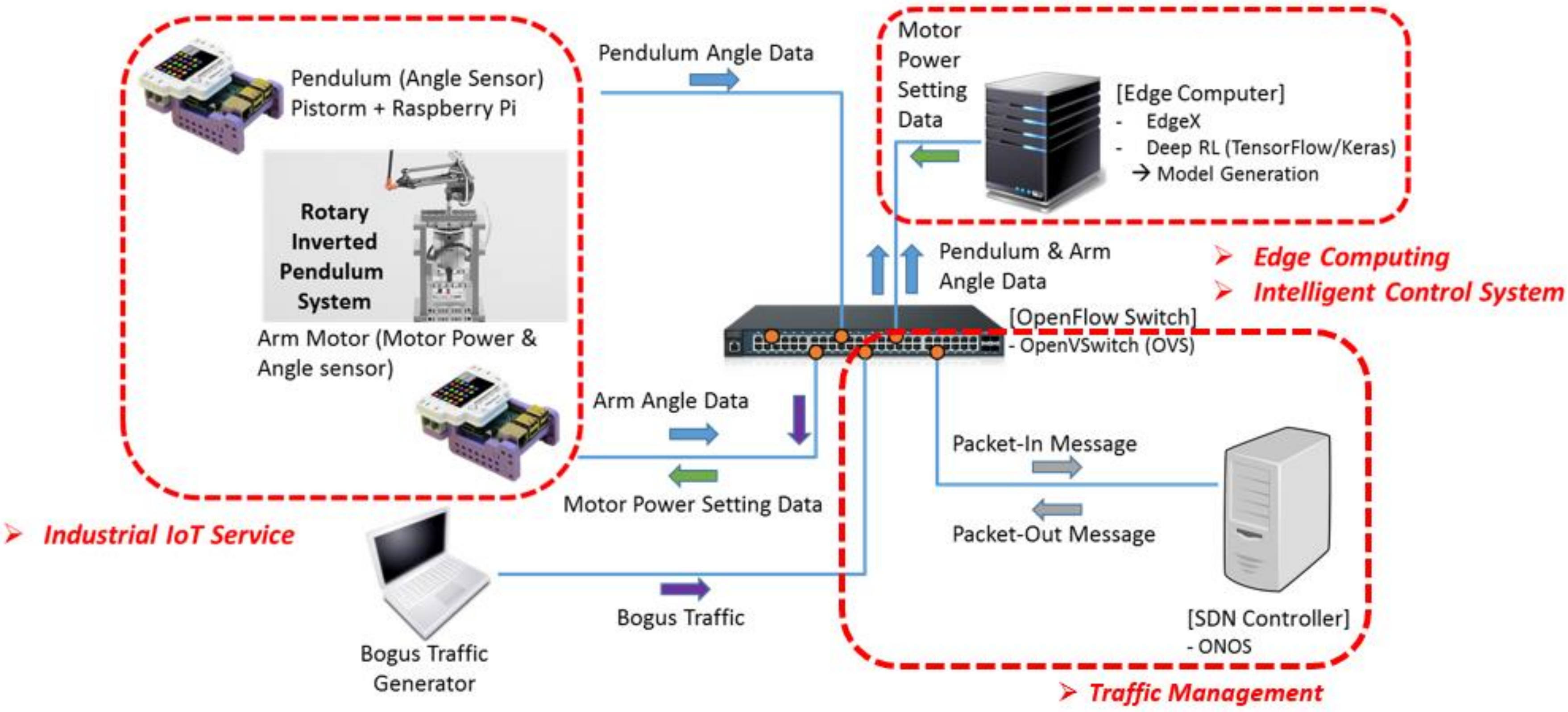
With Possible Network Scenario with RL

- **Intelligent Traffic Management**
  - Edge-based traffic management system for intelligent AI control using deep reinforcement learning to provide fast response time, reliability and security

- **Network Defect Prediction**
  - Deep reinforcement learning emerged the preferred solutions to manage and monitor the networking equipment (LTE core, router and switch) prevented by the networking failure risk

- **Routing Enhancement**
  - Routers in the multicast routing protocol are determined to discover optimal route with a predicted reward, and then the routers create the optimal path with multicast transmissions to reduce the overhead

Our Research:
Edge-based Intelligent Traffic Control system using DRL



Step1: DRL Simulation for Edge-based Control system

- **Simulation Environment**
  - OpenAI Gym (open source)
    - A toolkit for developing and comparing reinforcement learning algorithms (Environment)
  - DRL Definition
    - State, Action and Reward
    - Step, Episode
- **Improvement of Performance (Simulation)**
  - Deep-Q Network, A2C (Advantage-Actor-Critic), A3C (Asynchronous + A2C)

Simulation Environment for the System

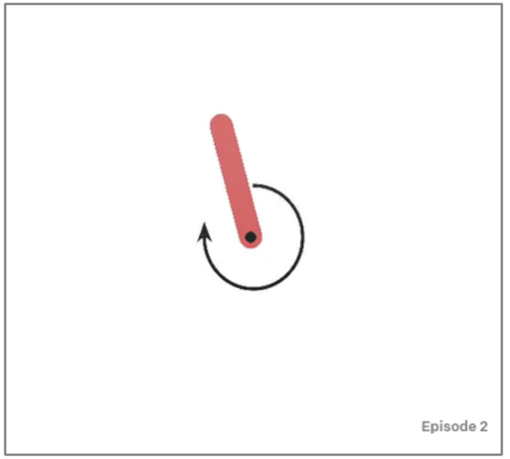
◆ Pendulum Environment (Pendulum v0)

- Observation

Num	Observation	Min	Max
0	$\cos\theta$	-1.0	1.0
1	$\sin\theta$	-1.0	1.0
2	$\theta'(=\frac{d\theta}{dt})$	-8.0	8.0

- Action

Num	Action	Min	Max
0	Joint effort	-2.0	2.0



Simulation Result (render())

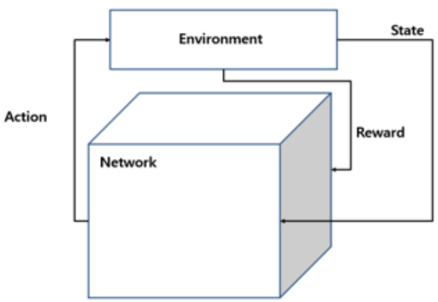
- Reward (per step)

$-(\theta^2 + 0.1 * \theta'^2 + 0.001 * action^2) = -(\pi^2 + 0.1 * 8^2 + 0.001 * 2^2) = -16.2736044$ → (** lowest cost)

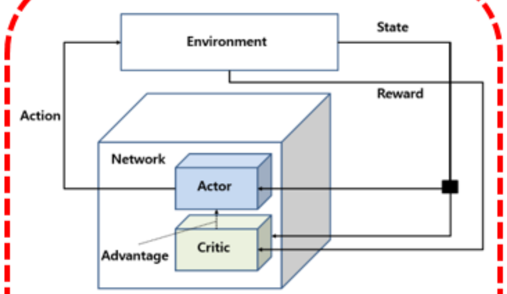
** highest cost = 0

Pendulum v0 · openai/gym Wiki · GitHub

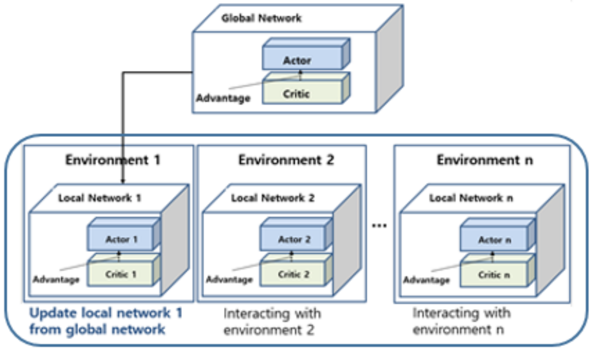
Simulation progress for Improvement of Performance



Deep-Q Network

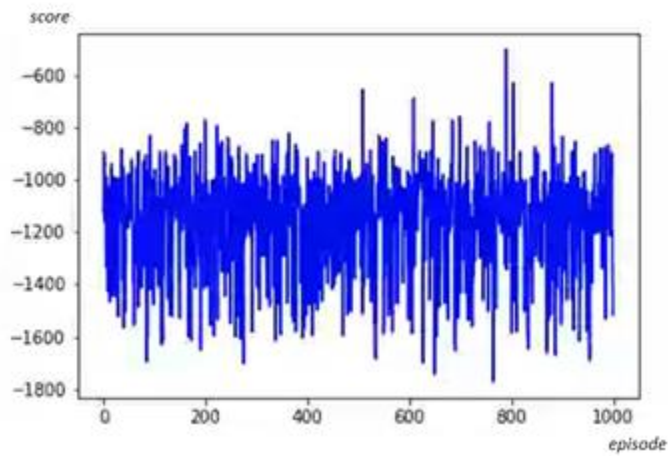


A2C

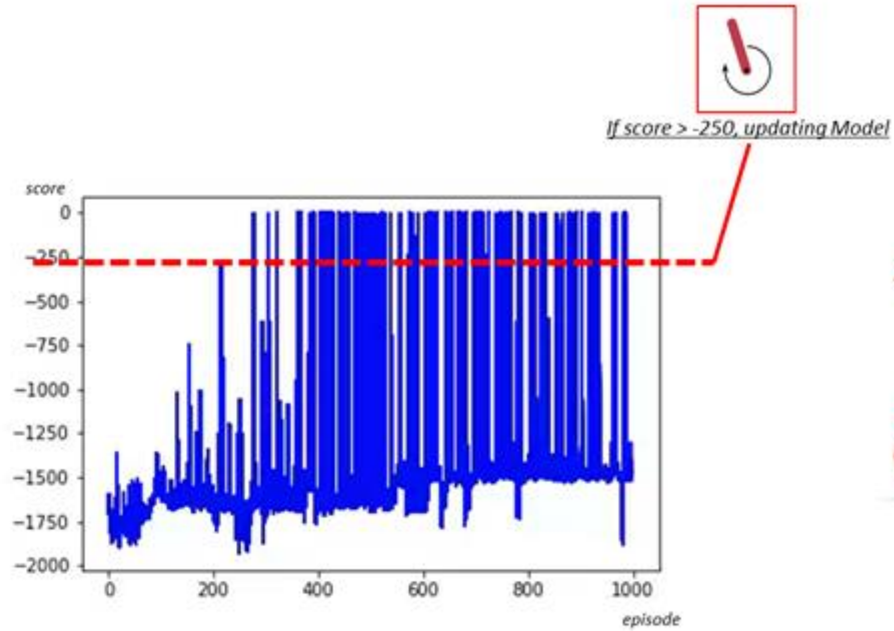


A3C

Simulation progress for Improvement of Performance



Deep-Q Network



A2C

A3C

More successful results than DQN

Next Steps for Contribution of NMRG :

- Adaptive open-flow network to intelligently manage and control traffic data
- Research and Implementation based on an edge-based platform
- AI engine in Edge-based platform using intelligent machine learning approach (DRL)

Thank you

Comment or Question?

Email:

mskim16@etri.re.kr

yghong@etri.re.kr