Deep Reinforcement Learning Untuk Pembelajaran Robot

Chandra Kusuma Dewa, Ph.D.
Jurusan Informatika, Universitas Islam Indonesia

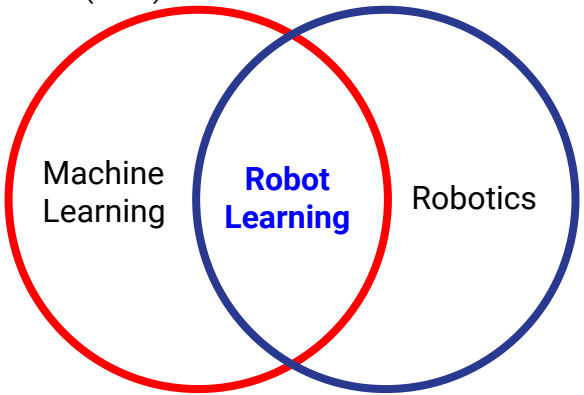
Disampaikan Pada Webinar Mingguan,
Indonesian Association for Pattern Recognition (INAPR)
Jumat, 8 April 2022

Biodata Pemateri

- **Nama**: Chandra Kusuma Dewa
- **Email**: chandra.kusuma@uii.ac.id
- **Latar Belakang Pendidikan**
  - **Ph.D.**, Toyohashi University of Technology, Japan (Oktober 2018 - September 2021)
  - **M.Cs.**, Universitas Gadjah Mada, Yogyakarta (Agustus 2008 - Februari 2011)
  - **S.Kom.**, Universitas Gadjah Mada, Yogyakarta (Agustus 2004 - Februari 2008)
- **Pengalaman Kerja**
  - 2013 - Sekarang: Dosen Informatika, Universitas Islam Indonesia
  - 2019: Machine Learning Engineer Intern, Honda R&D Co., Ltd., Saitama, Japan
- **Minat Penelitian**
  - Machine Learning
  - Reinforcement Learning
  - Robotics
  - Multimedia

Topik Penelitian Yang Ditekuni

Ibarz et al. (2021)

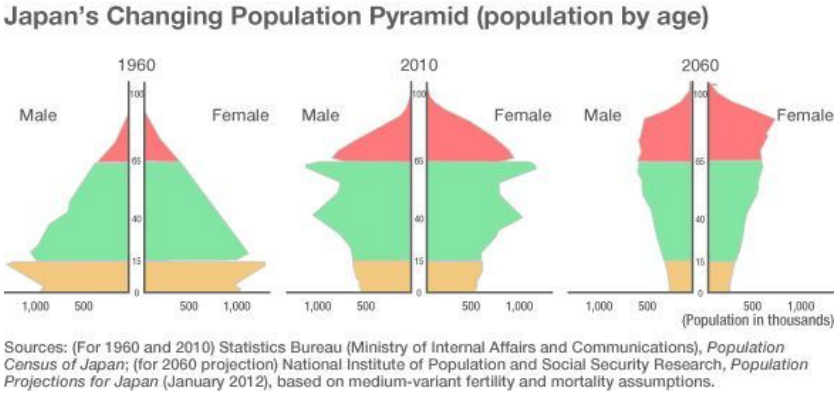


Gervasi et al. (2020)

Human-Robot Collaboration

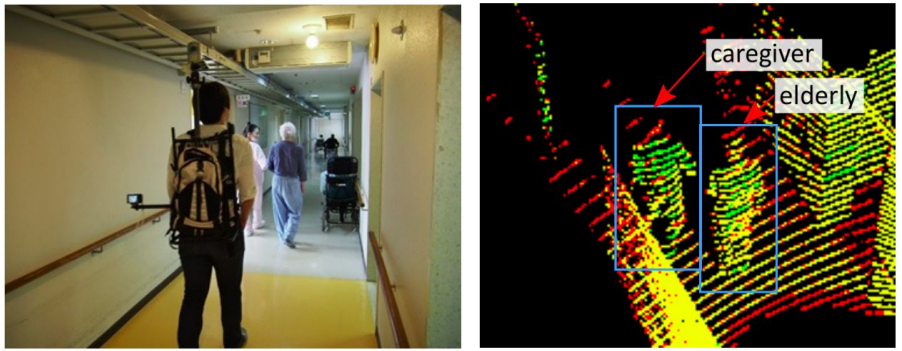
Attendant Robot Development
(One of AISL-TUT Research Topics)

Kenapa Jepang Mengembangkan Attendant Robot?



https://www.nippon.com/en/ncommon/contents/in-depth/41813/41813.jpg

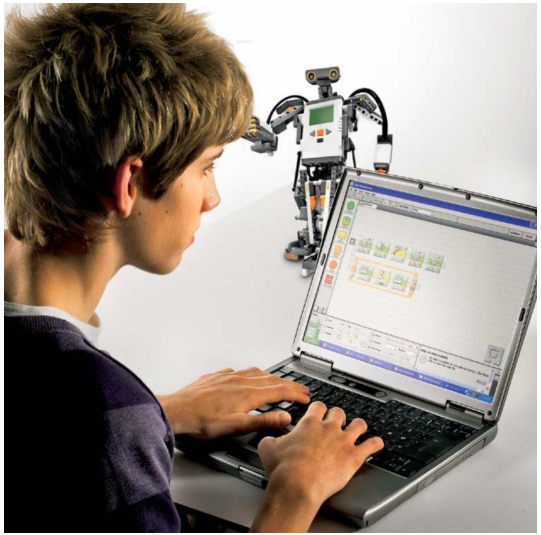
(Koide et al., 2019)

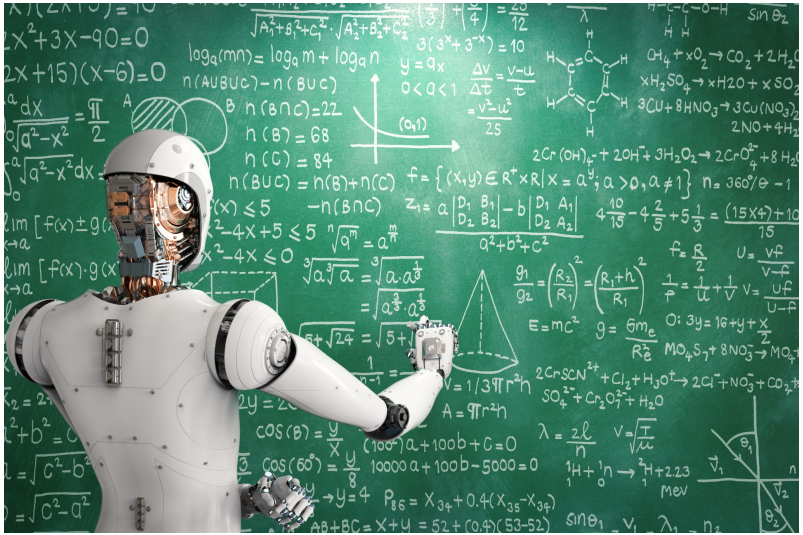


(a) Image. **(b)** Range data.

Sumber Gambar: https://www.nippon.com/en/ncommon/contents/in-depth/41813/41813.jpg

Bagaimana Mengembangkan Attendant Robot?





Directly program the robot

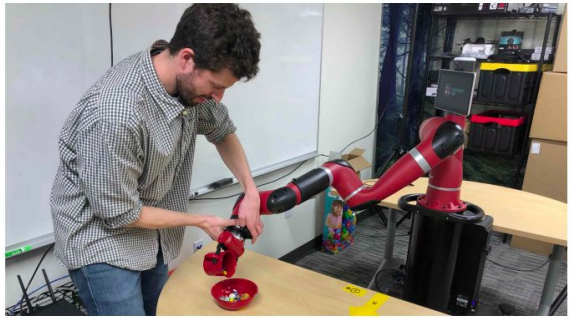
Can we make the robot able to learn by itself?

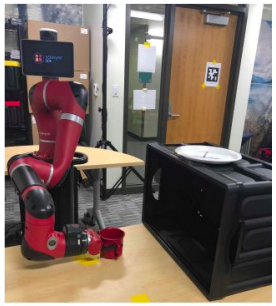
Is it possible? How?

Robot Learning

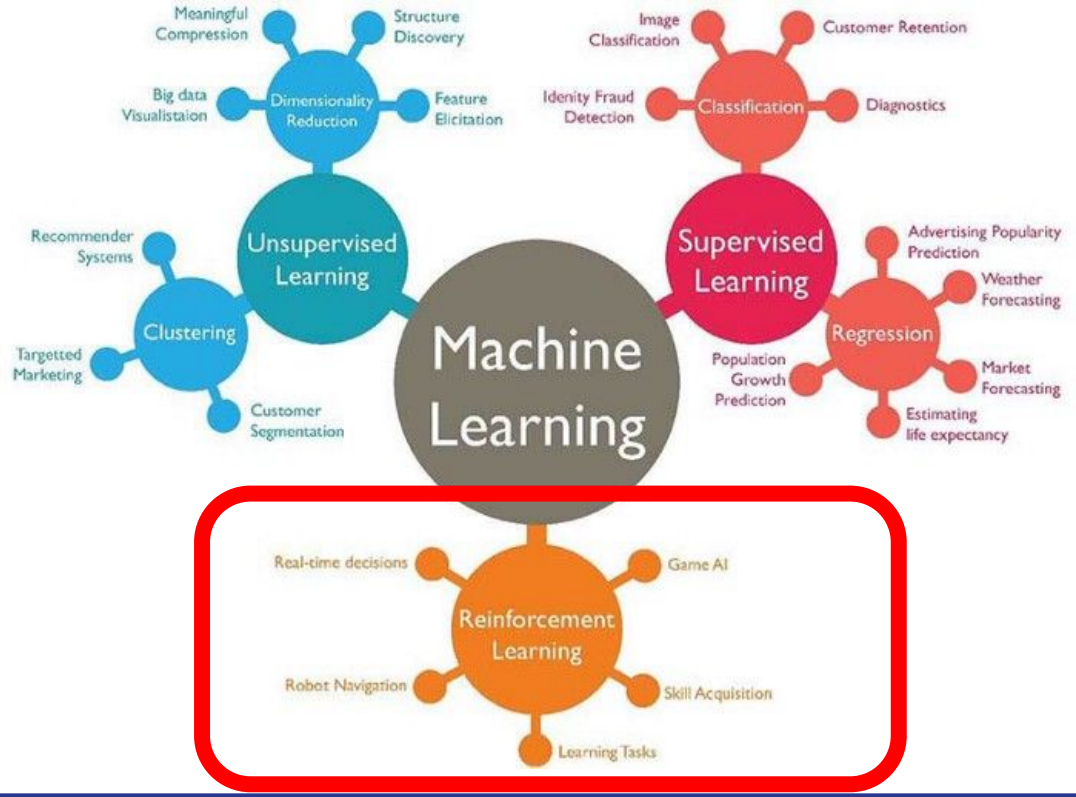
Left image source: https://www.robotshop.com/community/tutorials/show/how-to-make-a-robot-lesson-10-programming-your-robot
Right image source: https://bdtechtalks.com/wp-content/uploads/2019/04/artificial-intelligence-robot-math.jpg

Pembelajaran Robot dan Reinforcement Learning (1)





(Mueller et al., 2018)



Sumber Gambar Kanan:: https://miro.medium.com/max/700/1*iWJWMeOQmI5kmPMGZ3tovA.jpeg

Reinforcement Learning dan Pembelajaran Robot (2)

Sensors:

- **Camera**
- **Lidar**
- **IMU**
- **etc.**

DATA

(State)

Actions

- Move Forward
- Move Arm
- Stop
- Turn Left
- etc

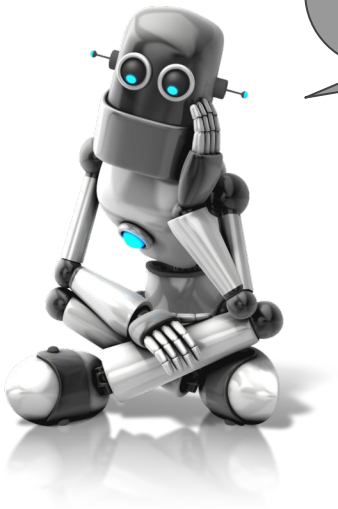
Robot (Agent)

Image Source: https://dej5y3fg919353ekq44ygvof-wpengine.netdna-ssl.com/wp-content/uploads/2021/03/Fetch-Mobile-Manipulator-Robot.jpg

Apa itu Reinforcement Learning?

According to Sutton and Barto (2017) Reinforcement Learning (RL) is a **goal-directed** learning from **interaction** that:

- Learns what to do
- Learns how to map situations to actions
- Learns to maximize a numerical reward signal



Mekanisme Pada Reinforcement Learning

I have to maximize the expected sum of reward

$$\sum_t \mathbb{E}_{(\mathbf{s}_t, \mathbf{a}_t) \sim \rho_\pi} [r(\mathbf{s}_t, \mathbf{a}_t)]$$

I will learn a **policy**

$$\pi(\mathbf{a}_t | \mathbf{s}_t)$$

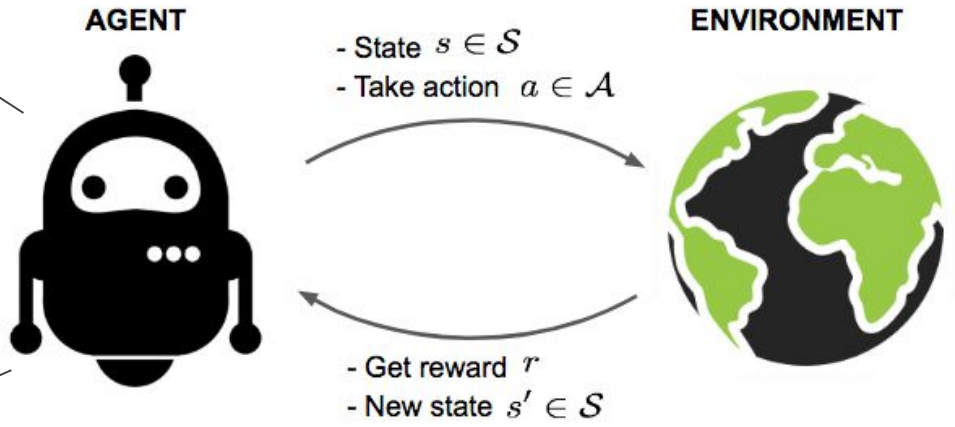
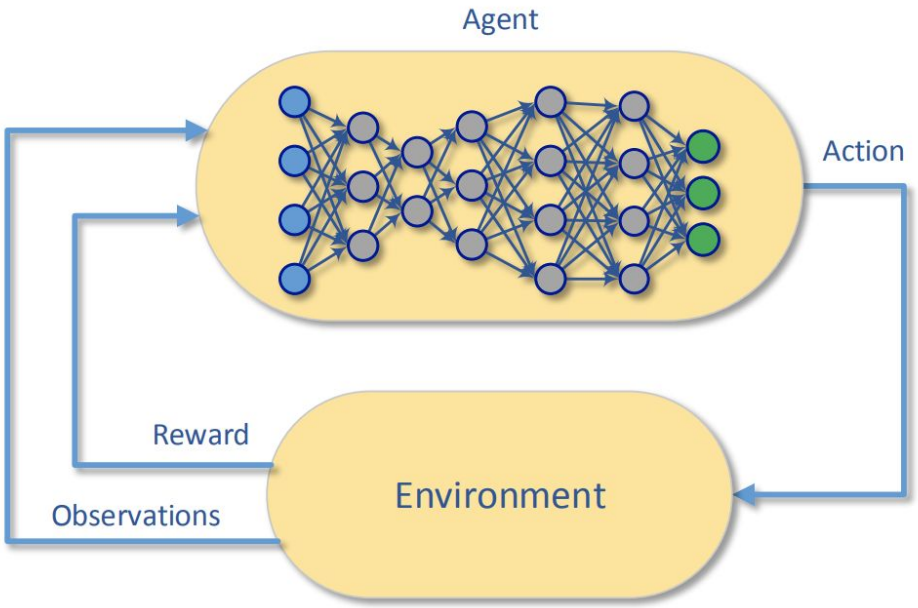


Image Source: https://lilianweng.github.io/lil-log/assets/images/RL_illustration.png

Reinforcement Learning Ingredients

- **Agent** - Implemented as a neural networks model which is designed to map states into actions
- **Environment** - Must follow Markov Decision Process (MDP) that:
  - Can be represented with set of states **S**
  - Accepts set of actions **A**
  - Has a state transition probability function **P**
  - Has a reward function **R**
- **Learning Algorithm** - Force the agent to maximize future rewards

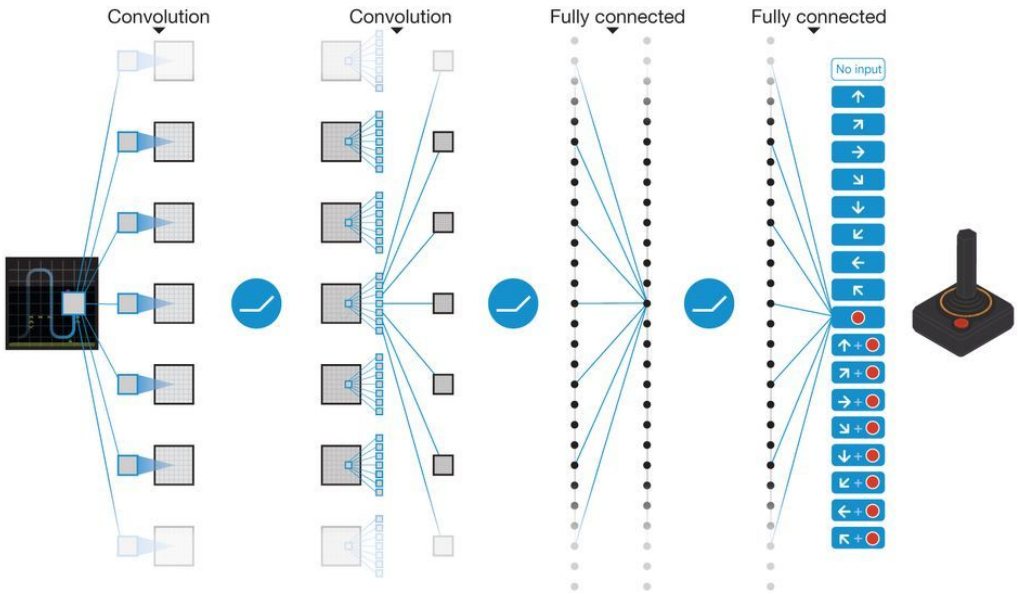
Deep Reinforcement Learning



- Deep Reinforcement Learning (DRL) is the combination of deep learning and reinforcement learning
- In DRL, the agent is implemented as **a deep learning model** so that it can map states into actions better

Image Source: https://arxiv.org/pdf/1810.04107.pdf

Contoh Implementasi DRL Pada Game



- In their paper in Nature, Mnih et al. (2015) implemented a DRL agent based on Convolutional Neural Networks (CNN)
- State in the environment is implemented as the video game screen

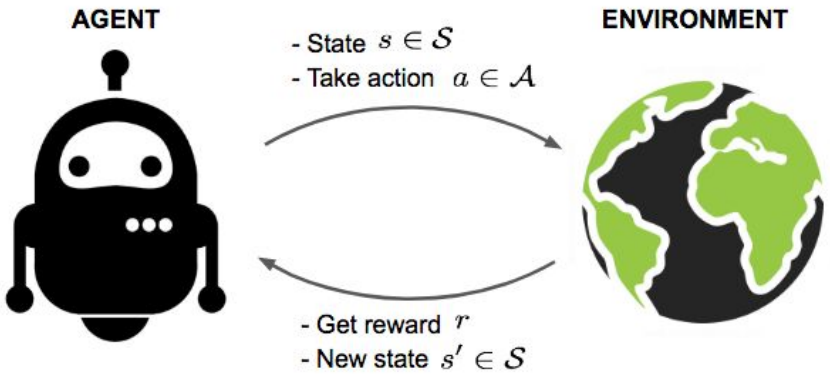
Image Source: https://www.datascienceassn.org/sites/default/files/Human-level%20Control%20Through%20Deep%20Reinforcement%20Learning.pdf

Contoh Proses Pembelajaran di Game

- **Environment** → Atari Breakout
- **Learning Algorithm** → Deep Q Learning
- **Reward Function** → Implemented based on the game's scores

Video source: https://www.youtube.com/watch?v=V1eYniJ0Rnk

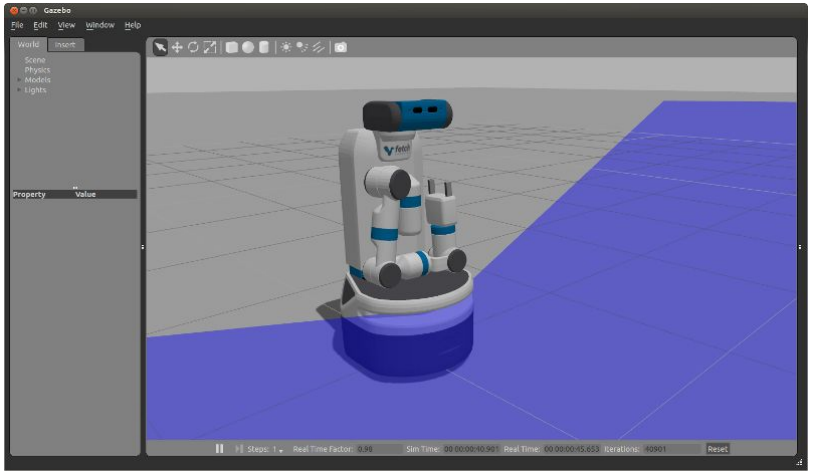
Implementasi DRL Untuk Pembelajaran Robot



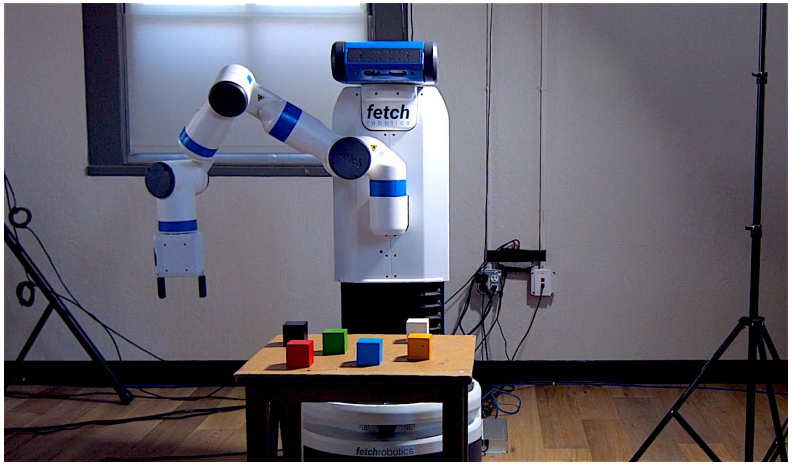
- Implementing the environments for the training process:
  - Realistic robot simulation
  - Real world environment
- Implementing the appropriate reward function which can guide the robot to perform a specific task correctly

Image Source: https://lilianweng.github.io/lil-log/assets/images/RL_illustration.png

Mengimplementasikan Environment



V.S.

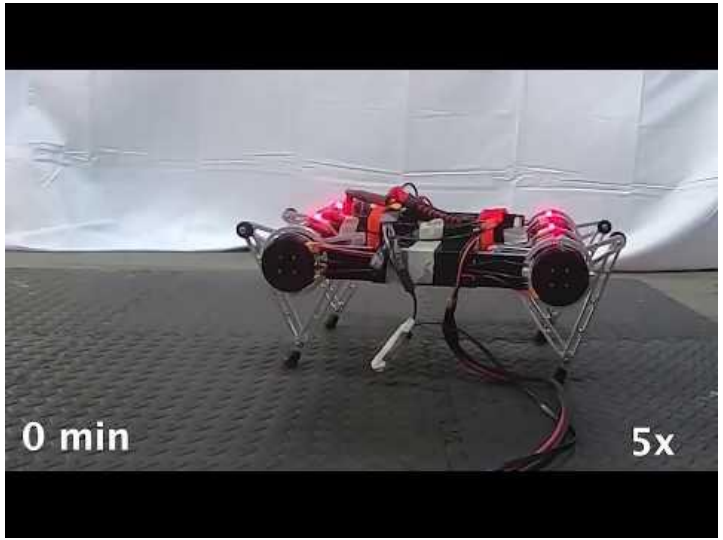


Simulation Environment (Gazebo simulator)

Real World Environment

Image Sources: https://docs.fetchrobotics.com/_images/gazebo.png, https://miro.medium.com/max/2840/1*HT-vGqmrcovVU1PDzcbw0A.png

DRL Training Pada Real World Environment



Haarnoja et al. (2019)

- Train a legged robot to walk directly in **real world**
- After 108 minutes, with Soft Actor-Critic learning algorithm, the robot can walk appropriately
- After obtaining the optimal policy, the robot even can walk properly under some disturbance

Video Source: https://www.youtube.com/watch?v=n2gE7n11h1Y&t=1s

DRL Training Pada Simulated Environment

- Train a bipedal robot to walk with **simulated environment** MuJoCo simulator)
- **Domain randomization** technique is used so that the optimal policy can be transferred from simulator to real world smoothly (**sim-to-real transfer**)

Li et al. (2021)

Video Source: https://www.youtube.com/watch?v=goxCjGPQH7U

Implementasi DRL Untuk Person-Following Robot

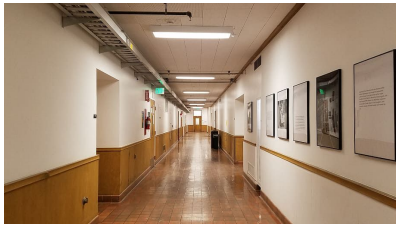
Pang et al. (2020)

reward = 1

reward = 1

reward = 1

reward = -10



Important tasks/aspects are excluded:

- Navigation
- Obstacle Avoidance
- Target person's comfortness

Video Source: https://www.youtube.com/watch?v=CuCXluONwkk

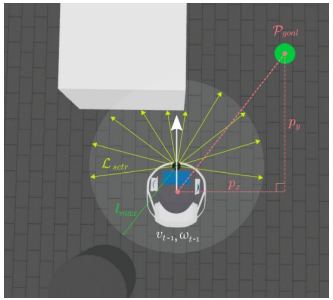
Person-Following Task Decomposition

Person-Following

Person-Following Policy

Navigation Task

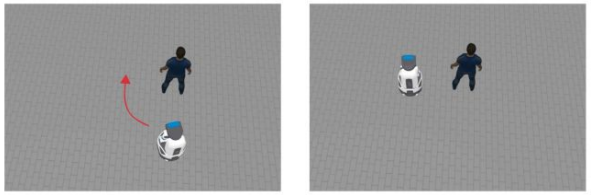
Navigation Policy



(Dewa & Miura, 2020)

Attending Task

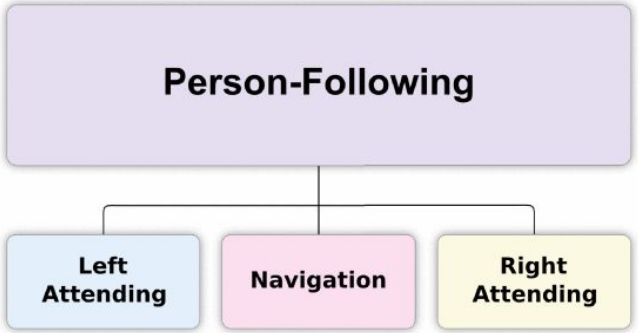
Attending Policy



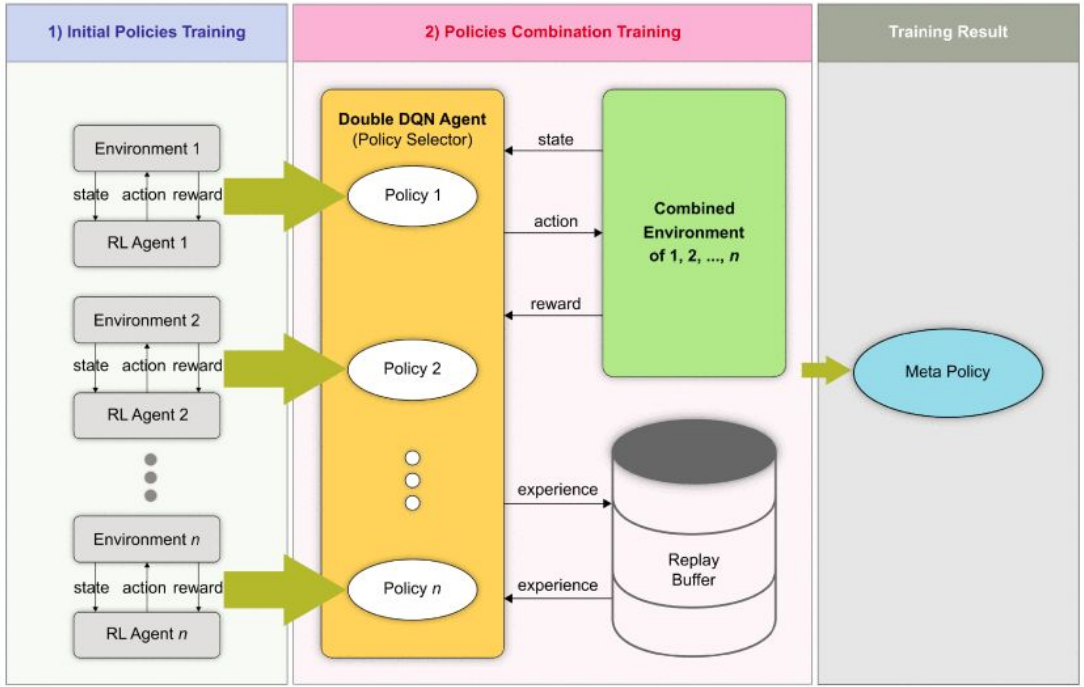
(Dewa & Miura, 2021)

Image Sources: A Framework for DRL Navigation With State Transition Checking and Velocity Increment Scheduling, IEEE Access (2020). Training a Robot to Attend a Person at Specific Location Using Soft Actor-Critic Under Simulated Environment, International Symposium of System Integration (SII) (2021)

Integrating The Navigation and The Attending Policies



(Dewa & Miura, 2021)



Terima Kasih