Autonomous Learning of Ball Trapping in the Four-legged Robot League

Hayato Kobayashi[1], Tsugutoyo Osaki[2],
Eric Williams[2], Akira Ishino[1],
Ayumi Shinohara[2]

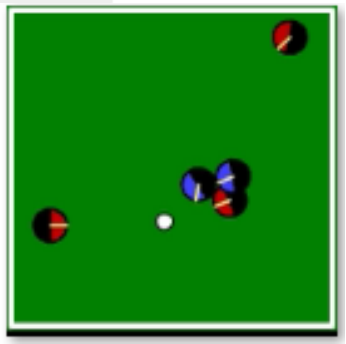
[1]Kyushu University, Japan
[2]Tohoku University, Japan

JollyPochie

Motivation

- Passwork in the four-legged robot league
  - KeepAway Soccer [Stone et al. 2001]
    - Benchmark of good passing abilities in the simulation league
  - Passing Challenge
    - Technical challenge in this year

It is too difficult for dogs



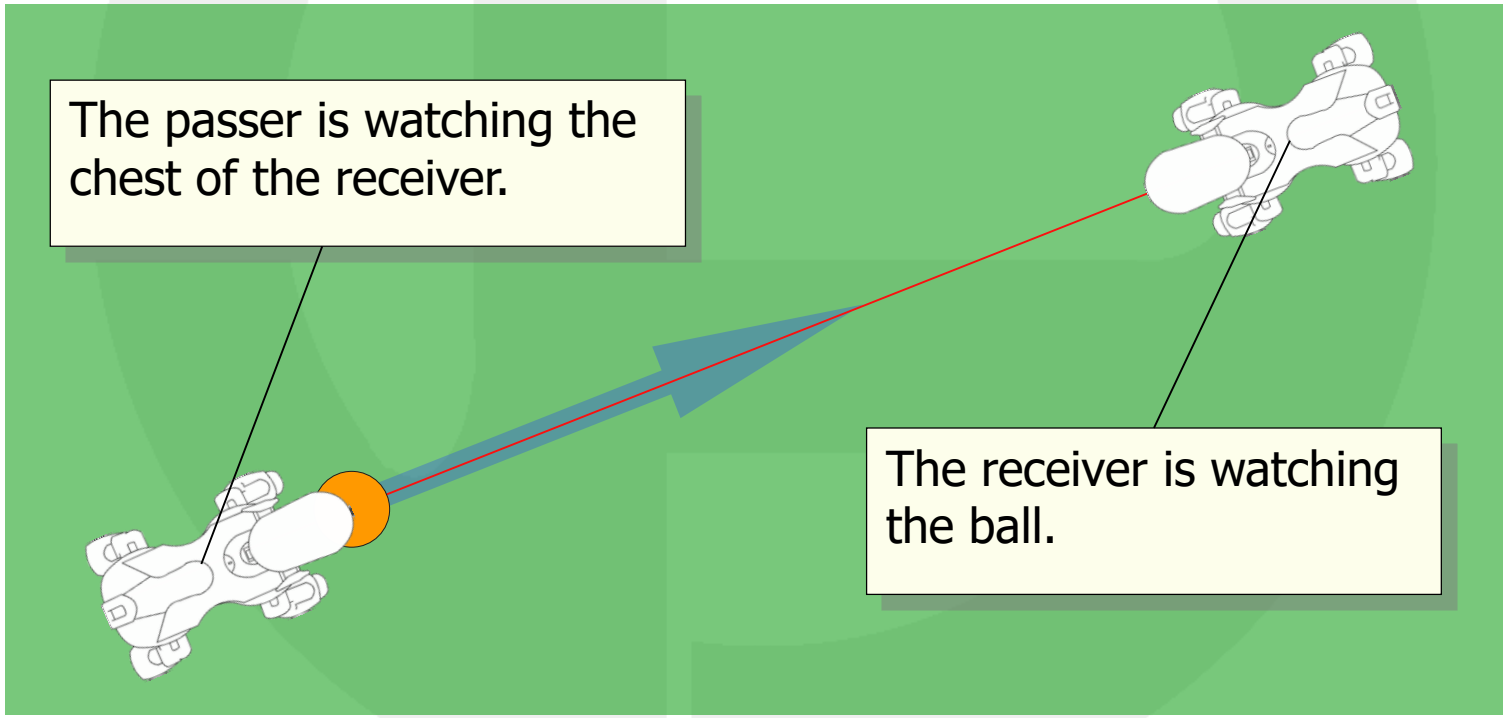
KeepAway Soccer
http://www.cs.utexas.edu/~AustinVilla/sim/keepaway/

JollyPochie

Ball Trapping

- Stop and control an oncoming ball

One-dimensional Model

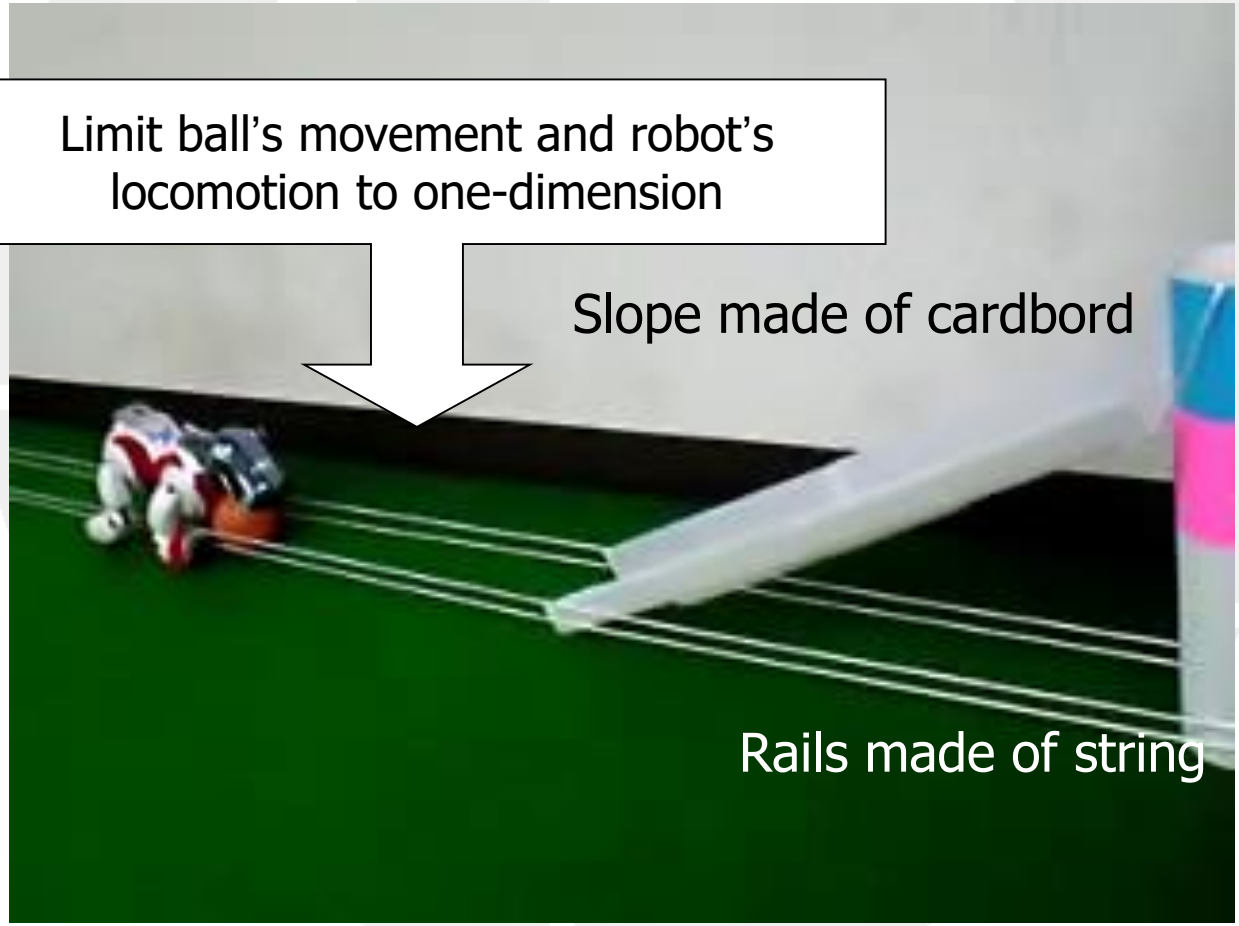


Autonomous Method

- Same way as diligent humans

Kick Wall

Training Equipment

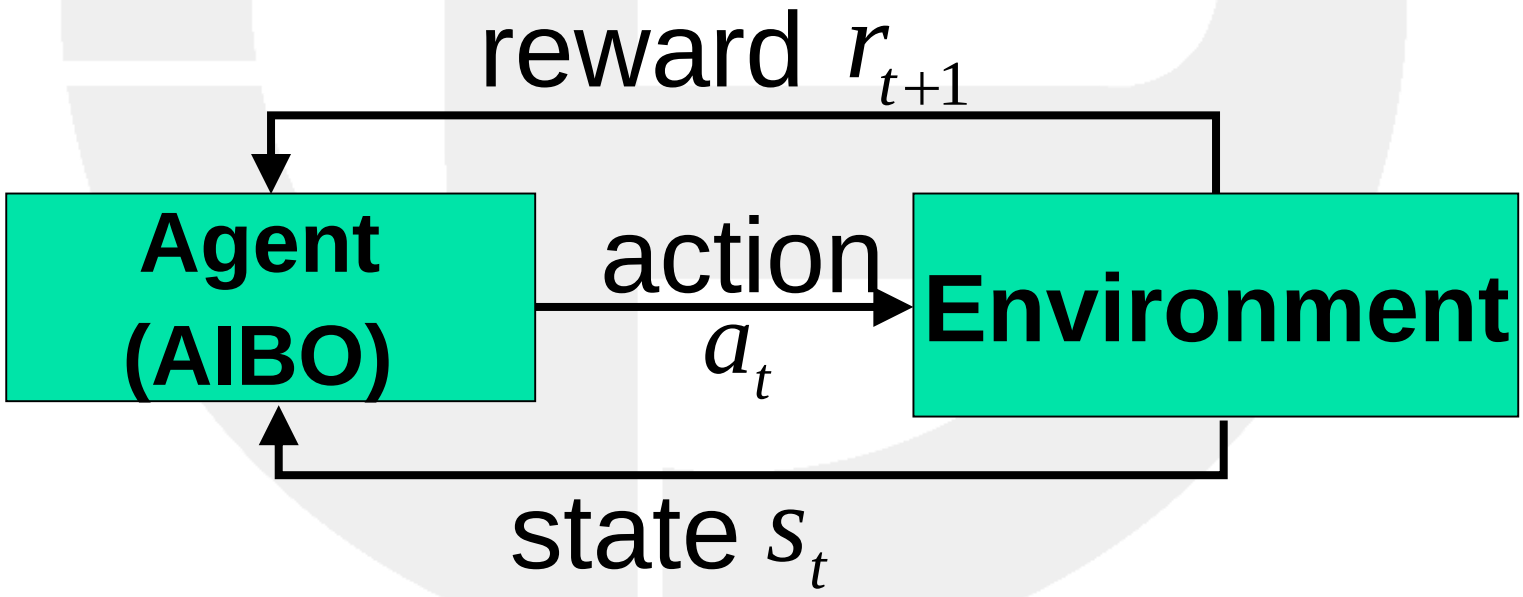


Learning Method

- Sarsa(λ) [Rummery and Niranjan 1994; Sutton 1996]
  - Reinforcement learning algorithm
- Tile-coding (aka CMACs [Albus 1975])
  - Linear function approximation
  - For speeding up their learning

Reinforcement Learning

- Acquire maps from state input to action output maximizing the sum of rewards



In our study, each time step t = 0, 1, 2, … mean 0ms, 40ms, 80ms, …

Implementation

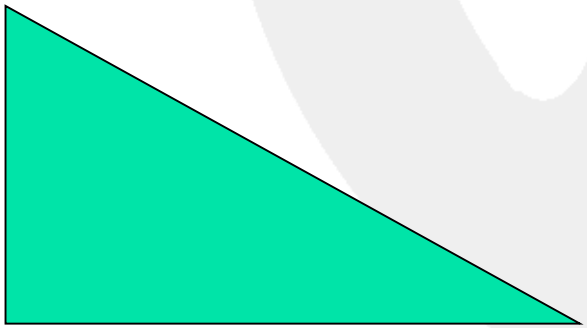
- State $s_t = (x_t, dx_t)$
  - x_t・・・The distance from the robot to the ball
    [0,2000]（mm）
  - dx_t・・・The difference between the current x_t and the previous x_t of one time step before.
    [-200,200]（mm）
- Action a_t
  - *ready*・・・Move its head to watch the ball
  - *trap*・・・Initiate the trapping motion

Implementation

- Reward r_{t+1}
  - Positive
    - If the ball was correctly captured between the chin and the chest after the *trap* action.
  - Negative
    - If the *trap* action failed, or
    - If the ball touches the chest PSD sensor before the *trap* action is performed.
  - Zero
    - Otherwise

Implementation

- Episode
  - The period from kicking the ball to receiving any reward other than zero

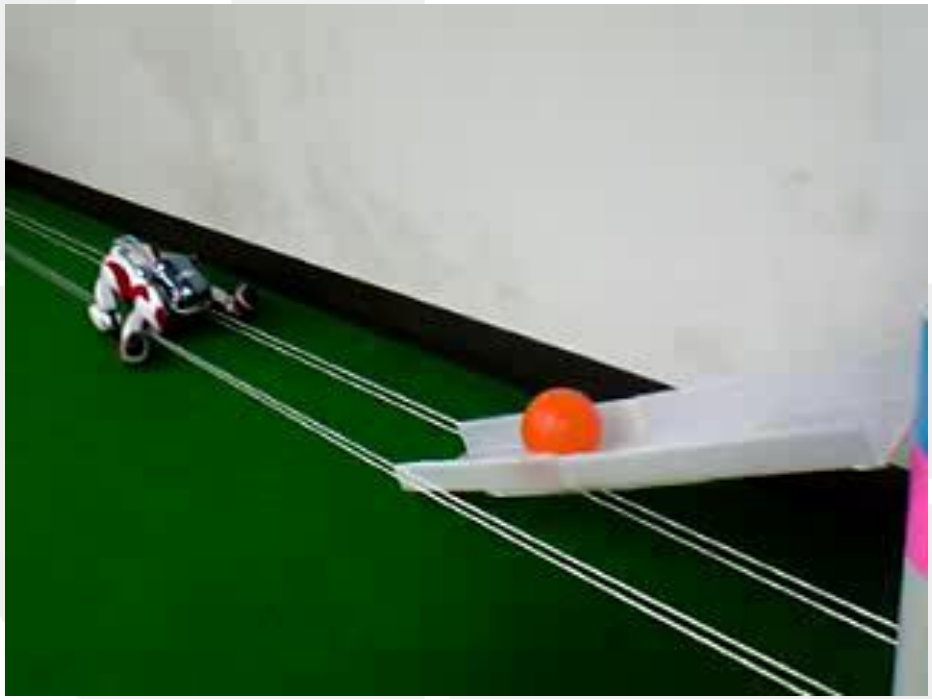


Experiments

- Using one robot
- Using two robots
  - without communication
  - with communication

Using One Robot

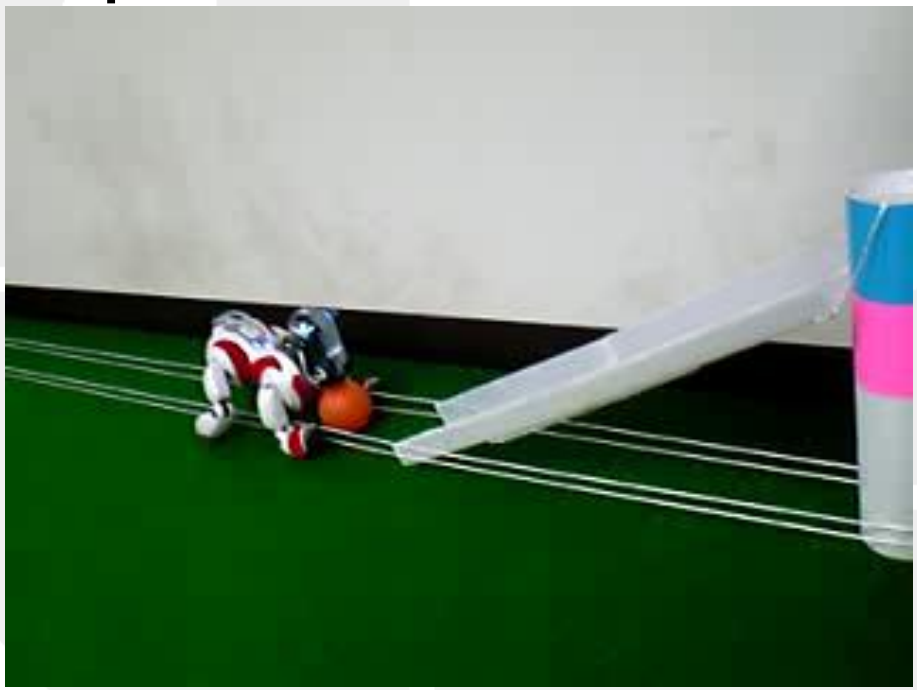
- Earlier phase



https://youtu.be/hv1sgIZLpKA

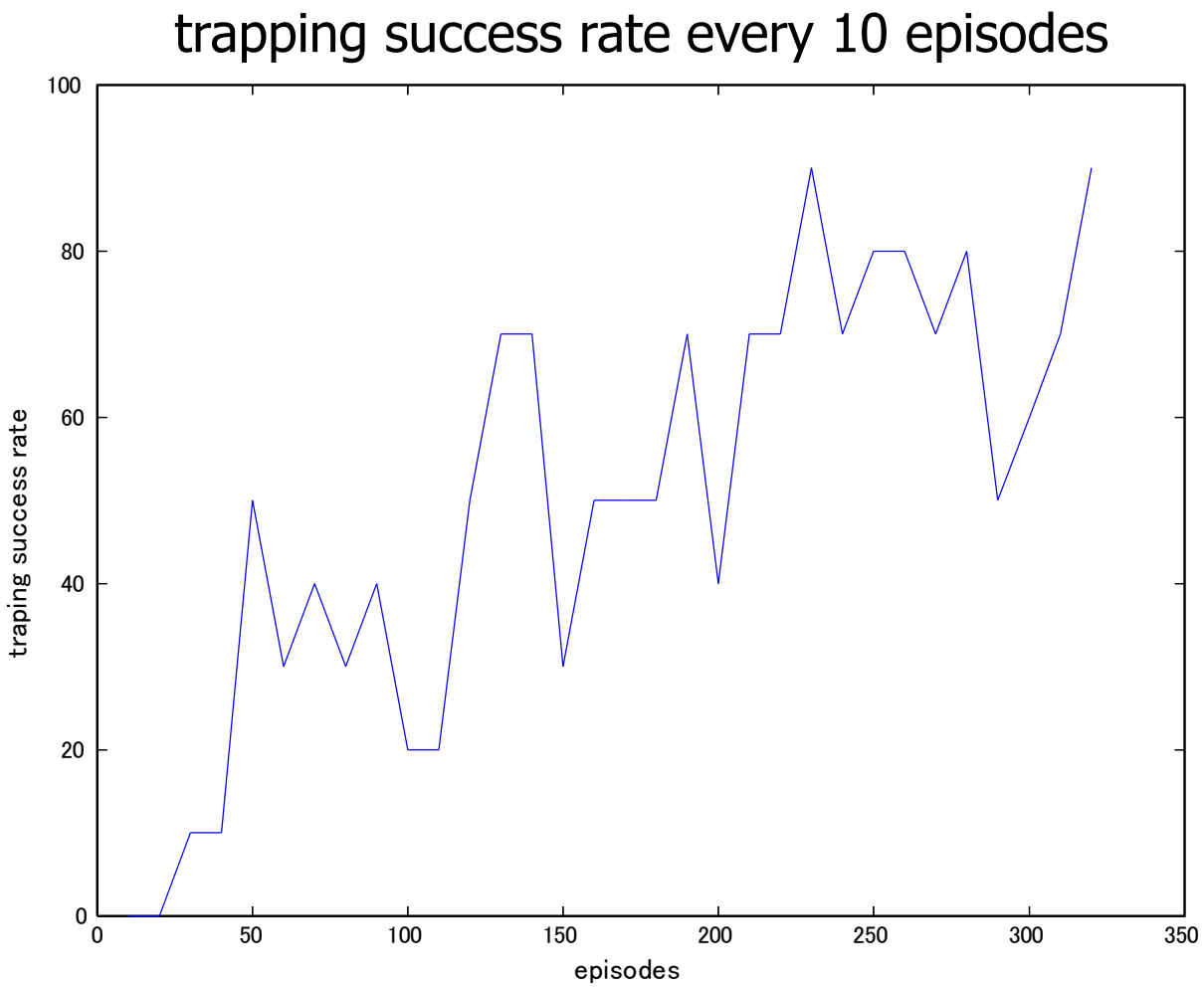
Using One Robot

- Later phase

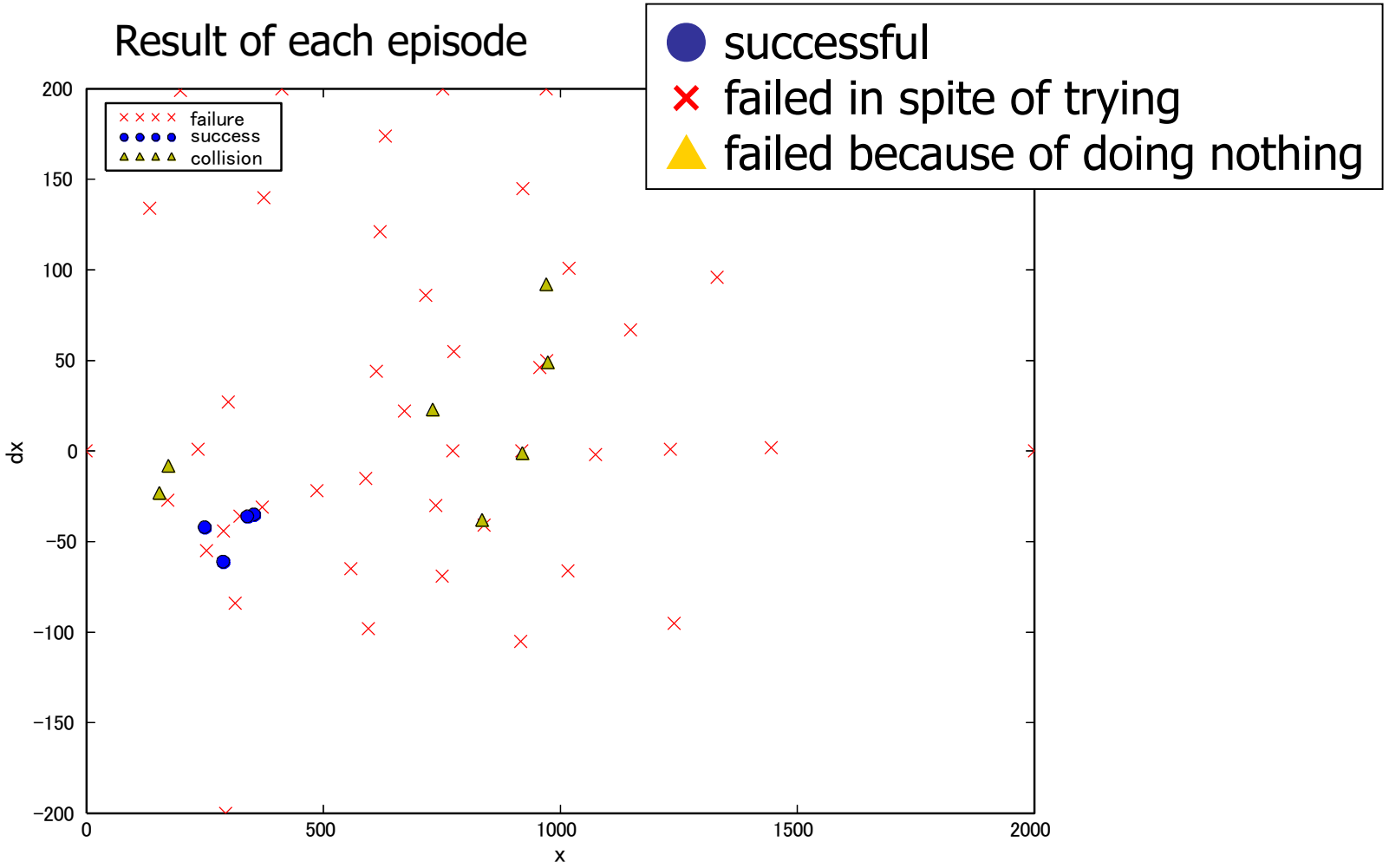


https://youtu.be/XJBllv7wJXQ

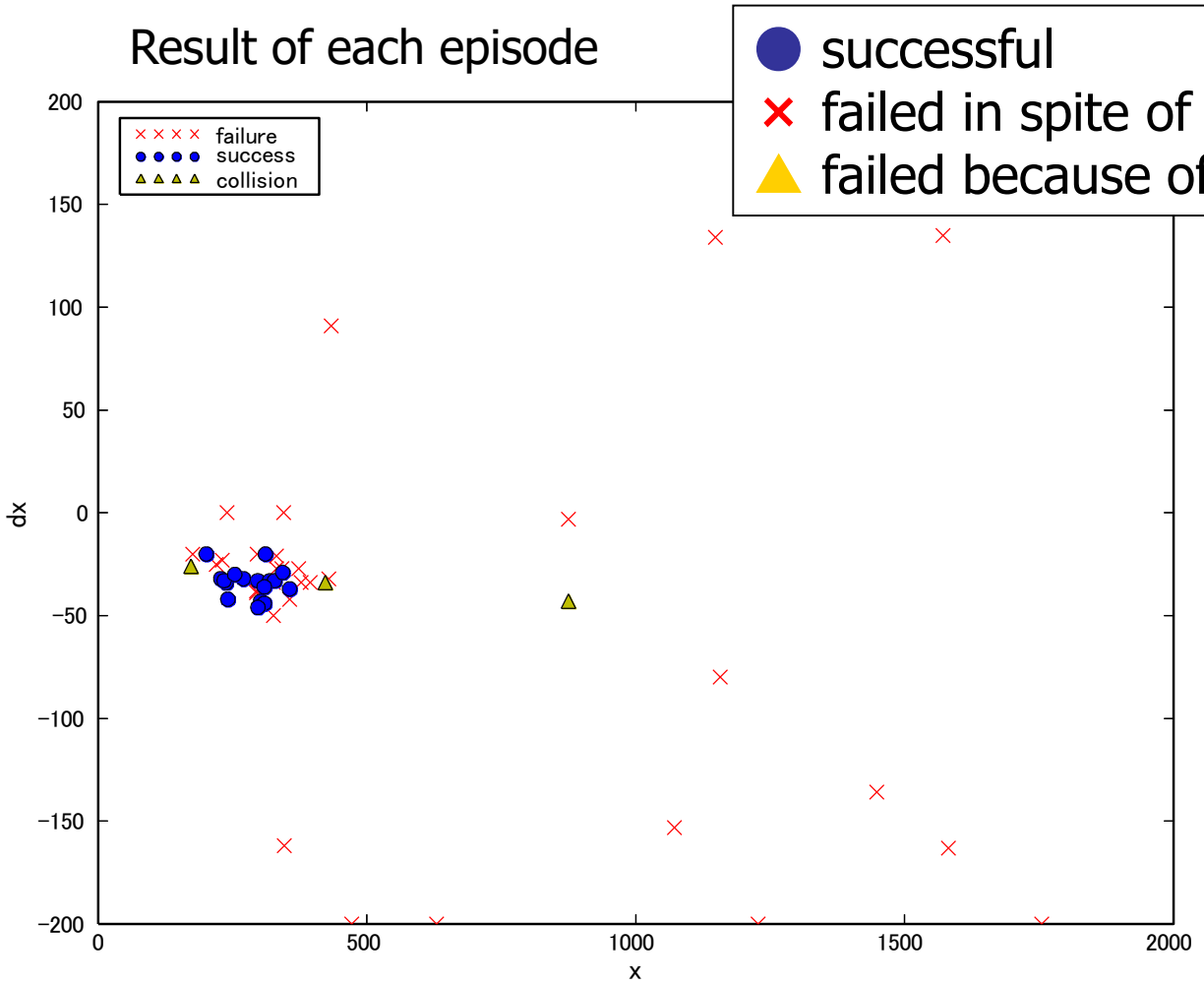
Result of Learning Using One Robot



Episodes 1…50

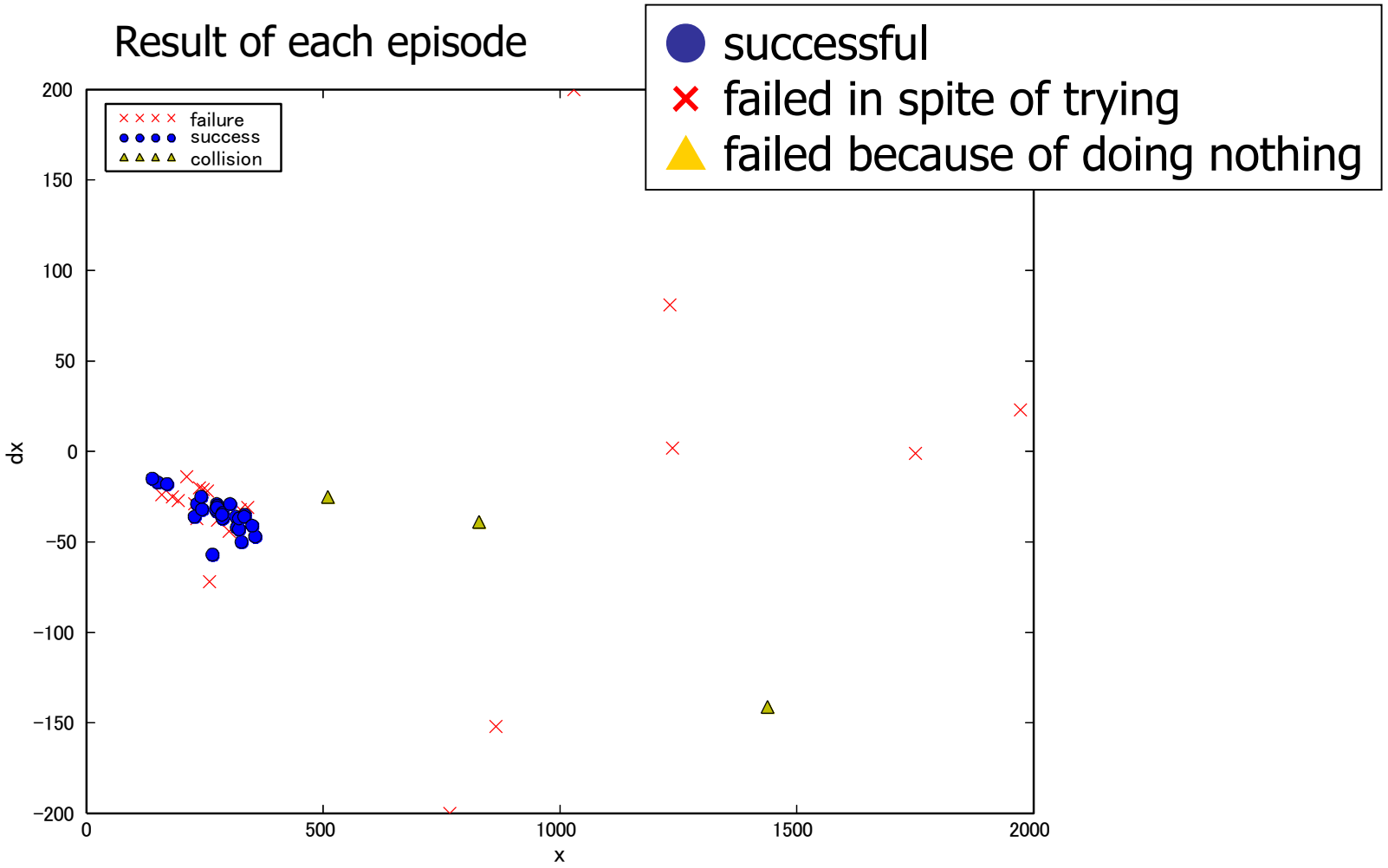


Episodes 51…100

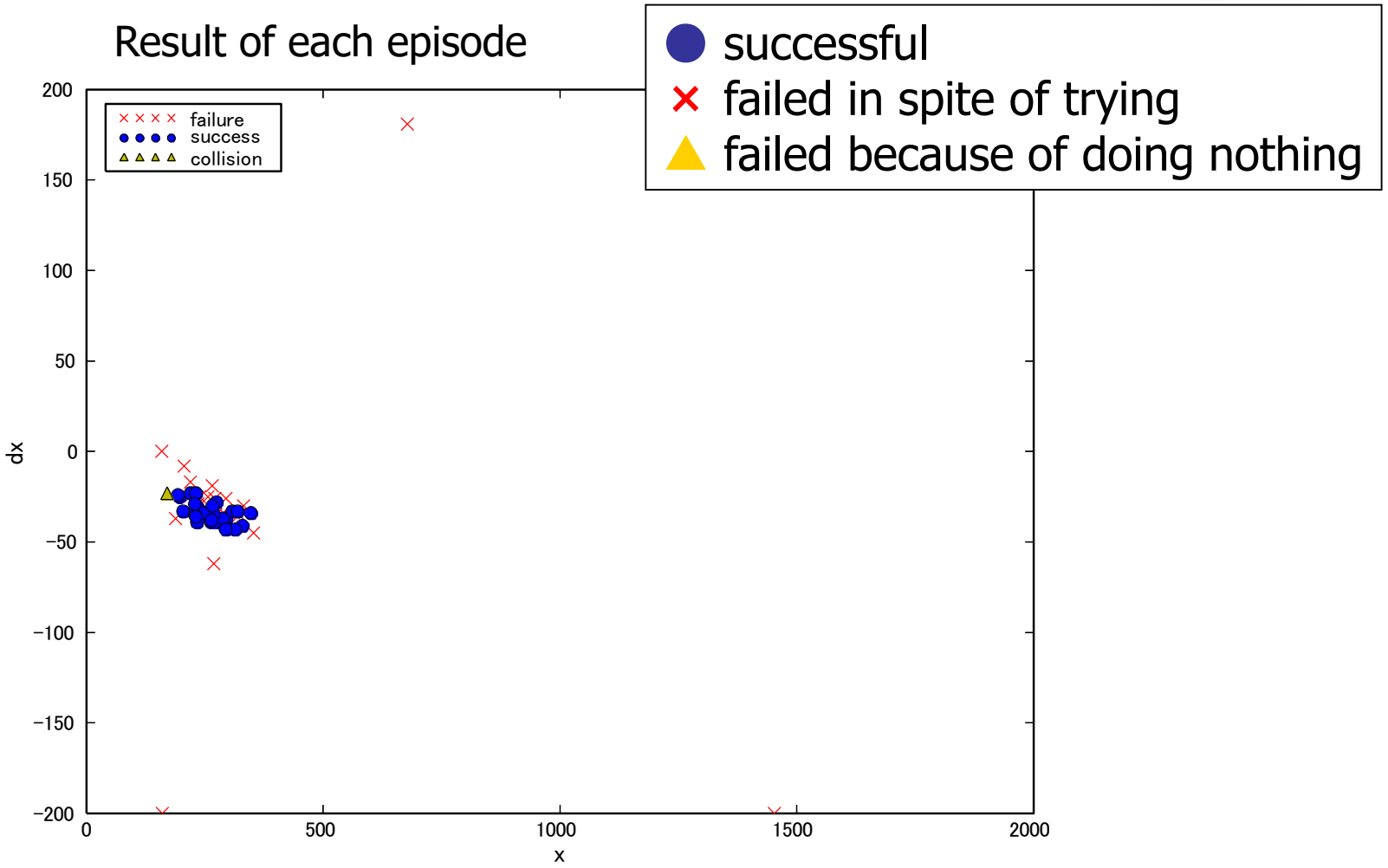


- successful
- failed in spite of trying
- failed because of doing nothing

Episodes 101…150



Episodes 151…200

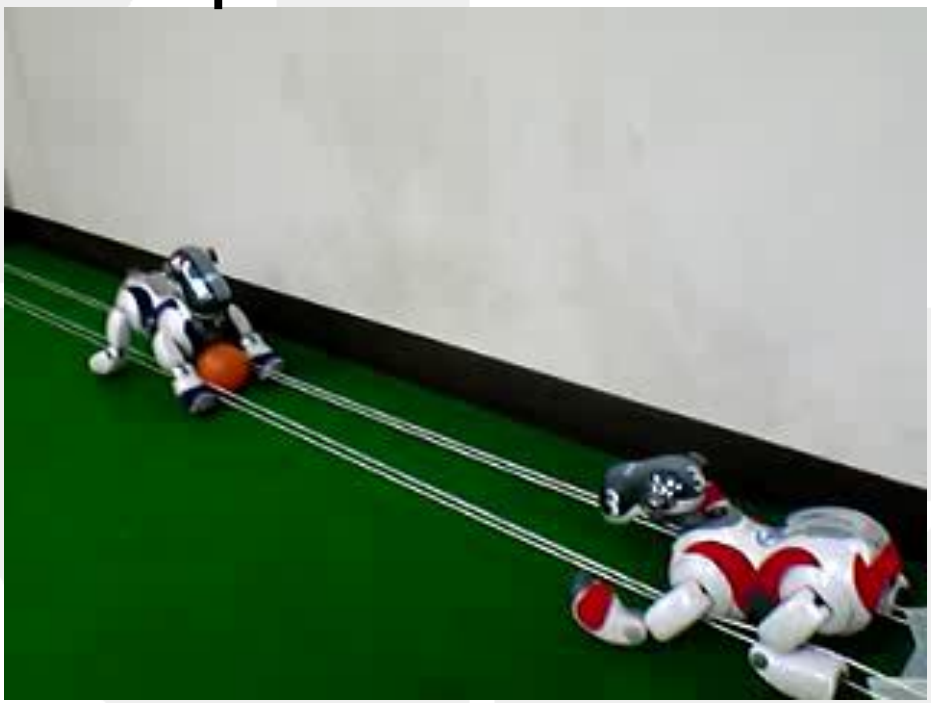


Using Two Robots

- Simply replace slope with another robot
  - Active Learner (AL)
    - Original robot
    - Same as in case of training using one robot
  - Passive Learner (PL)
    - Replaces slope
    - Does not approach the ball if the trapping failed

Using Two Robots

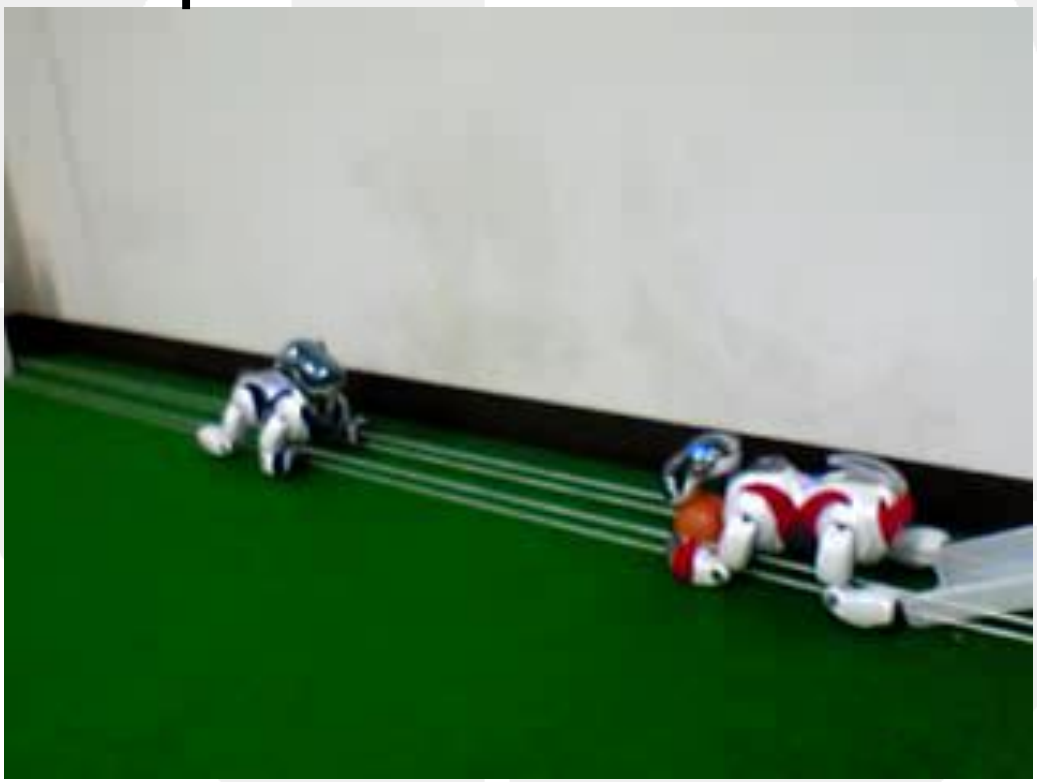
- Earlier phase



https://youtu.be/sXkVYZjOzjg

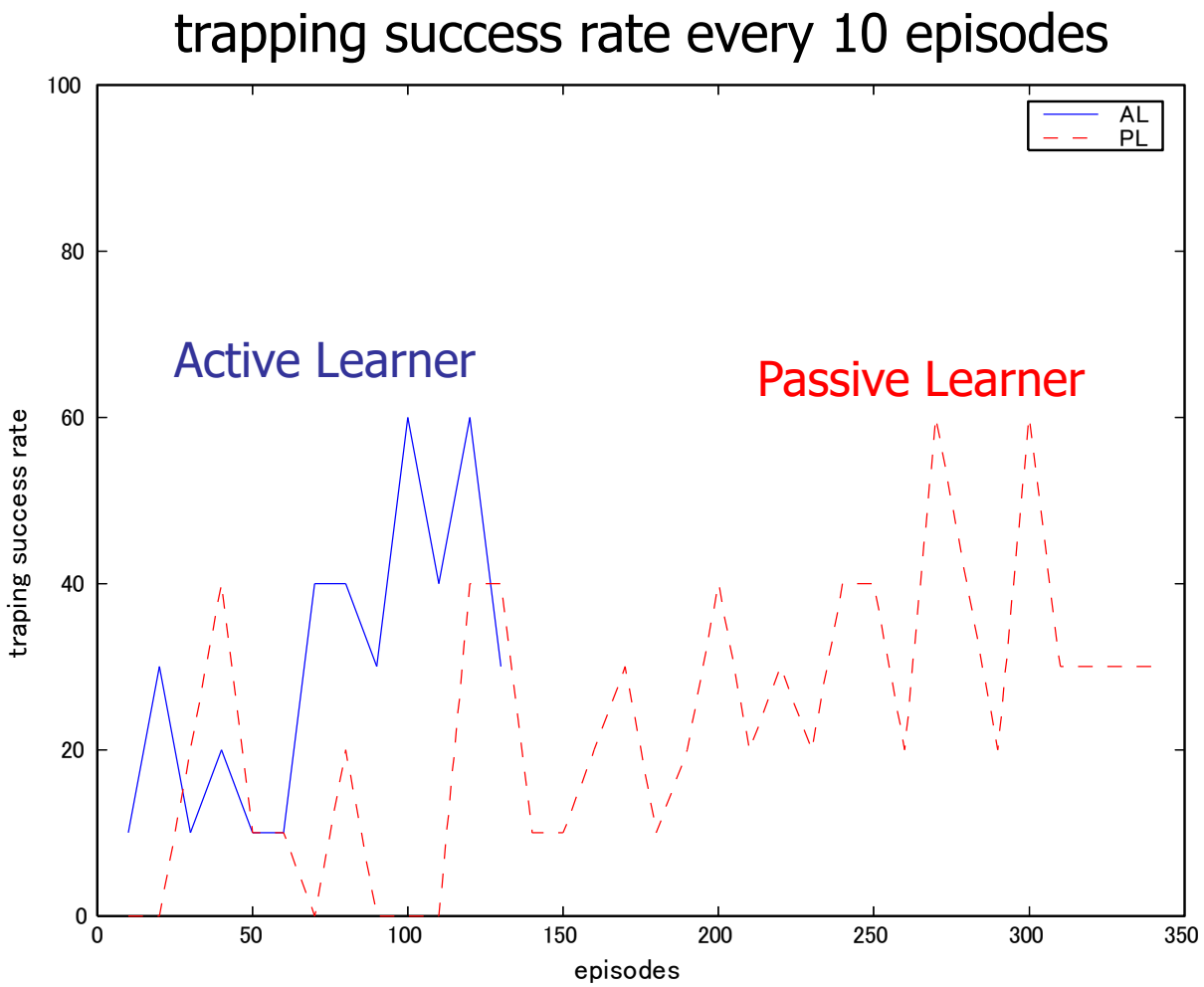
Using Two Robots

- Later phase



https://youtu.be/opvoyv9h-GU

Result of Learning Using Two Robots Without Communication



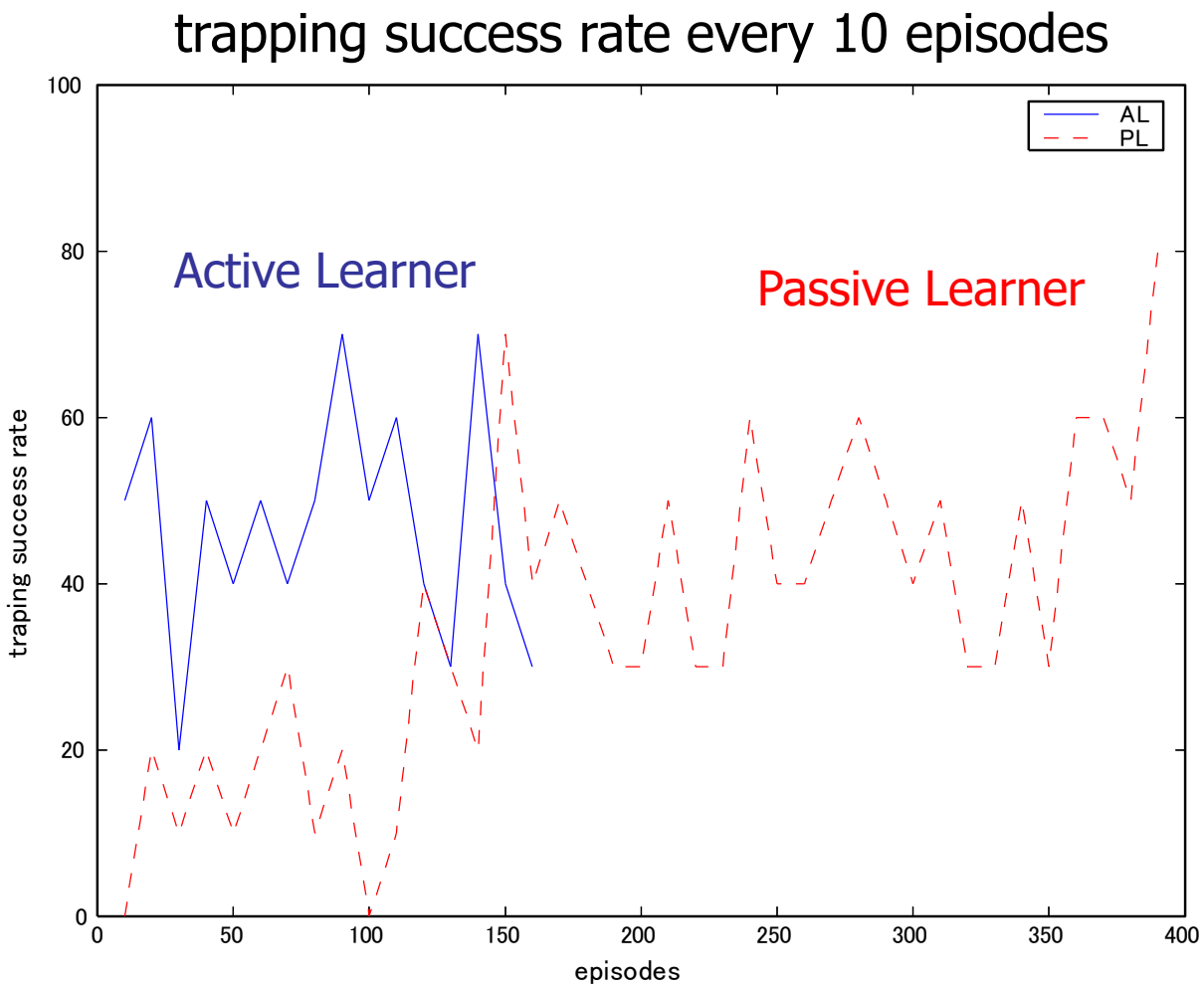
Problem of Using Two Robots

- Takes a long time to learn
  - AL can only learn when PL itself succeeds
    - Cannot learn if the ball is not returned
  - Even if we use only two ALs, the problem is not resolved
    - Just learn slowly, though simultaneously.

Solution

- Sharing their experiences
  - Their experiences include
    - Action a_t (*trap* or *ready*)
    - State variables $s_t = (x_t, dx_t)$
    - Reward r_{t+1}

Result of Learning Using Two Robots With Communication



Conclusion

- The goal of pass-work is achieved in one-dimension
  - learned the skills without human intervention
  - learned more quickly by exchanging experiences with each other

Future Work

- Extend trapping skills to two-dimensions
  - Layered Learning [Stone 2000]
  - Make goalies stronger
- Make robots learn passing skills simultaneously

Thank you for your attention!

Bremen is a good town!

JollyRobbie