End-to-End Reinforcement Learning

Reinforcement Learning is learning only for action selection (RL) and learning for entire process by using with a (recurrent) neural network (RNN).

Our Standpoint

Recent Revolution in Image Recognition

Deep Learning

learning of parallel processing

1000 categories

Through learning, useful features emerge

The power of end-to-end learning has been proven. Recognition is not just classification but sometimes like thinking!

It is impossible to decide all such recognition targets in advance. The reason why we have to decide the recognition targets in advance is that we isolate the recognition module from the others!

* Extend to the entire process from sensors to motors
* Introduce Reinforcement Learning (RL) for autonomous learning
* Introduce Recurrent Neural Network (RNN) for learning of dynamics

Function Emergence through End-to-End RL

[Shibata, 1997] [Hassabis (Google DeepMind), 2016]

Interesting Contextual Behavior (2008)

Objective of this task

Go to switch

Flag signals on the switch

Flag on goal 1

Flag on goal 2

Actor-Critic (TD)

NN (BP)

Hidden representation

Q-learning (TD)

Networks have to discover the action meanings and learn to remember it.

Pattern Meaning Discovery Task (2011)

Objective of this task

Go to switch

Flag signals on the switch

Flag on goal 1

Flag on goal 2

Actor-Critic (TD)

NN (BP)

Hidden representation

Q-learning (TD)

Networks have to discover the action meanings and learn to remember it.

Comprehensive Prediction Task (2013)

Actor-Q (TD)

NN (BP)

Hidden representation

Q-learning (TD)

Networks have to discover the action meanings and learn to remember it.

In the next poster (No.21), you can find our works about the emergence of communication! A new proposal towards the emergence of "thinking" will be presented!