Alibaba Group
阿里巴巴集团

星际争霸与人工智能

阿里巴巴认知计算实验室 龙海涛

Alibaba Group
阿里巴巴集团
Why StarCraft?
0
0
120/65
Protoss Pylon
Psi Used: 120
Psi Provided: 8
Total Psi: 65
Psi Max: 200
300/300 300/300
MENU
Replay Progress
Elapsed Time: 00:06
Speed: Fast

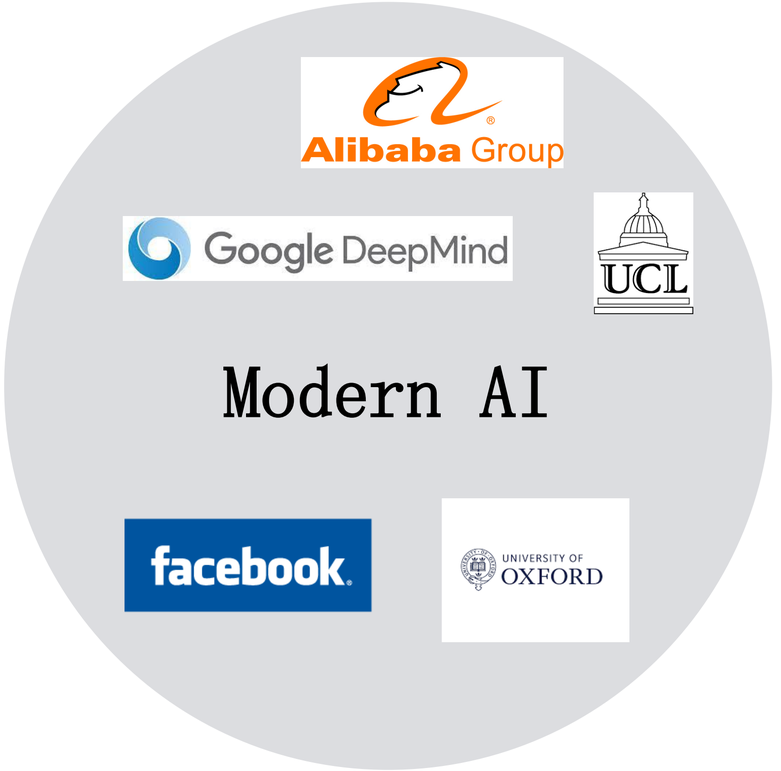
Alibaba Group 阿里巴巴集团

Challenge Problems for Artificial Intelligence

- Imperfect Information
- Huge State and Action Space
- Adversarial Real-time Strategy
- Long-Term Planning
- Temporal and Spatial Reasoning
- Multiagent Cooperation

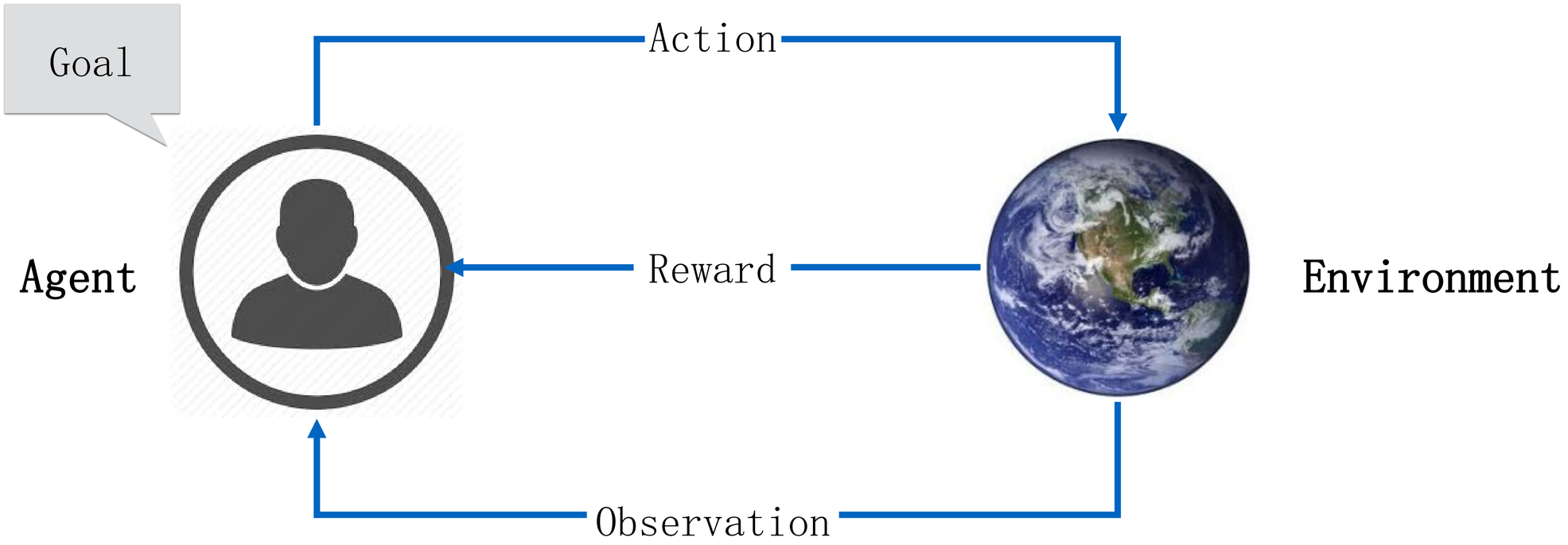
StarCraft AI Research and Competitions

2010~Now

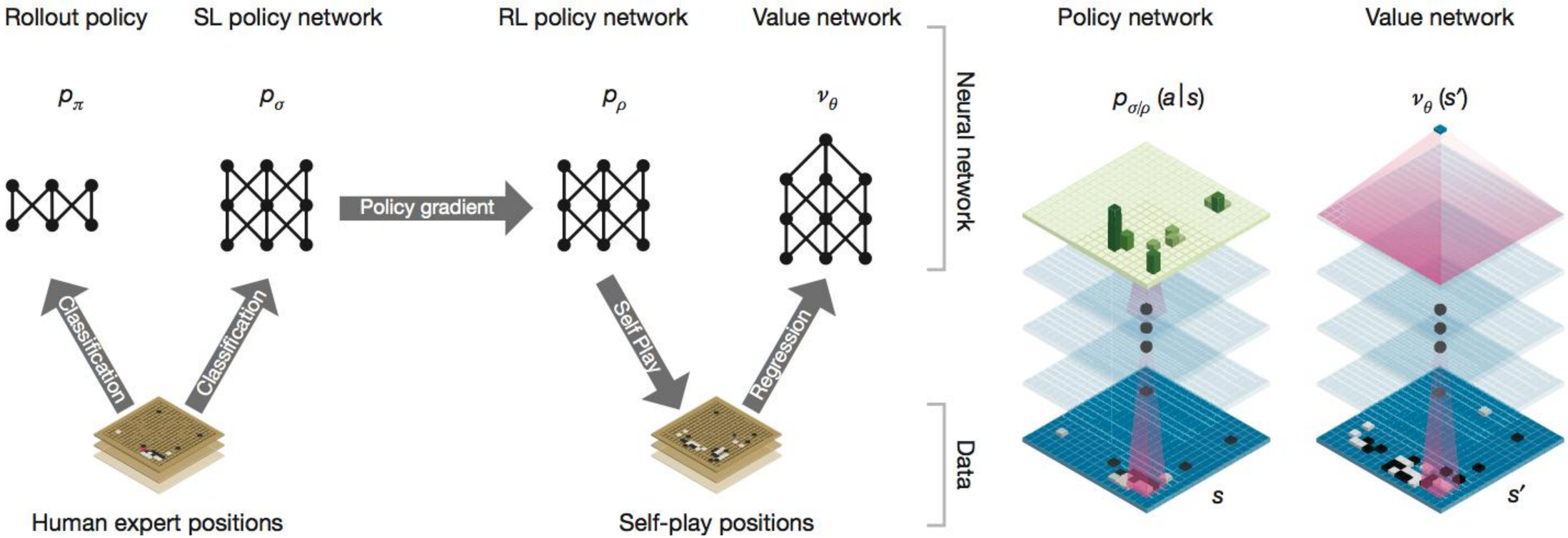


2016~Now

Reinforcement Learning



Deep Reinforcement Learning

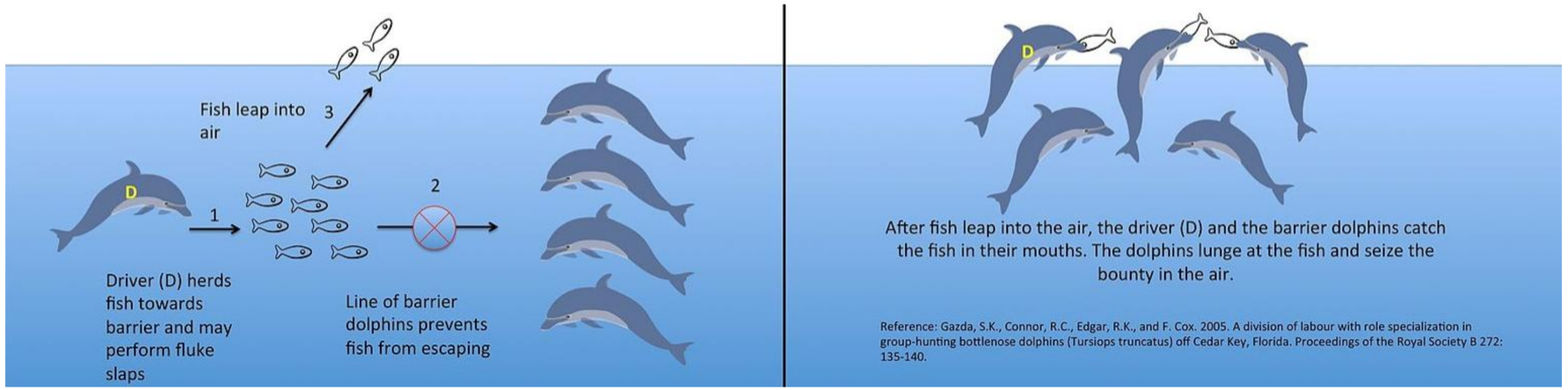


Alibaba Group
阿里巴巴集团

What is next?

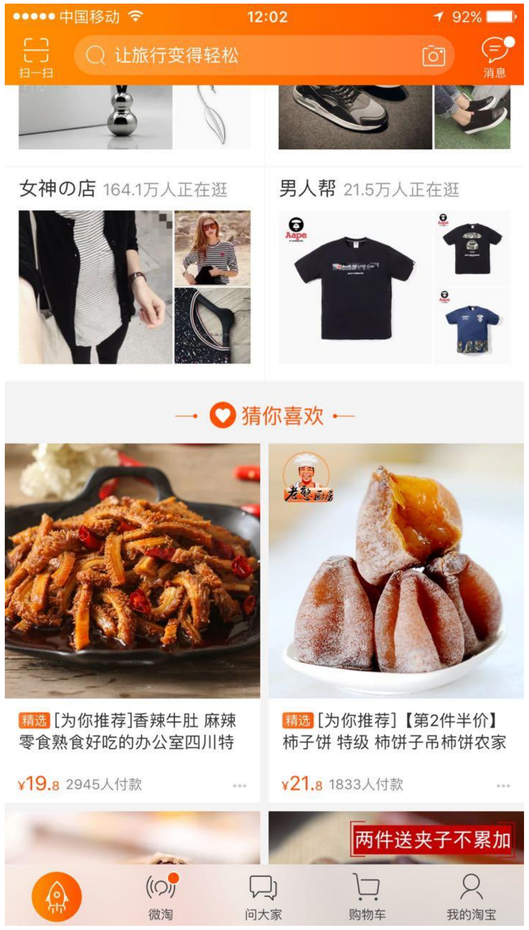
- All above are single AI agent
- But, true human intelligence embraces social and collective wisdom
- How large-scale multiple AI agents could learn human-level collaborations (or competitions) from their experiences?

Cooperative Hunting



Alibaba Group
阿里巴巴集团

Artificial Collective Intelligence



Alibaba Group 阿里巴巴集团

Multiagent Bidirectionally-Coordinated Nets for Learning to Play StarCraft Combat Games

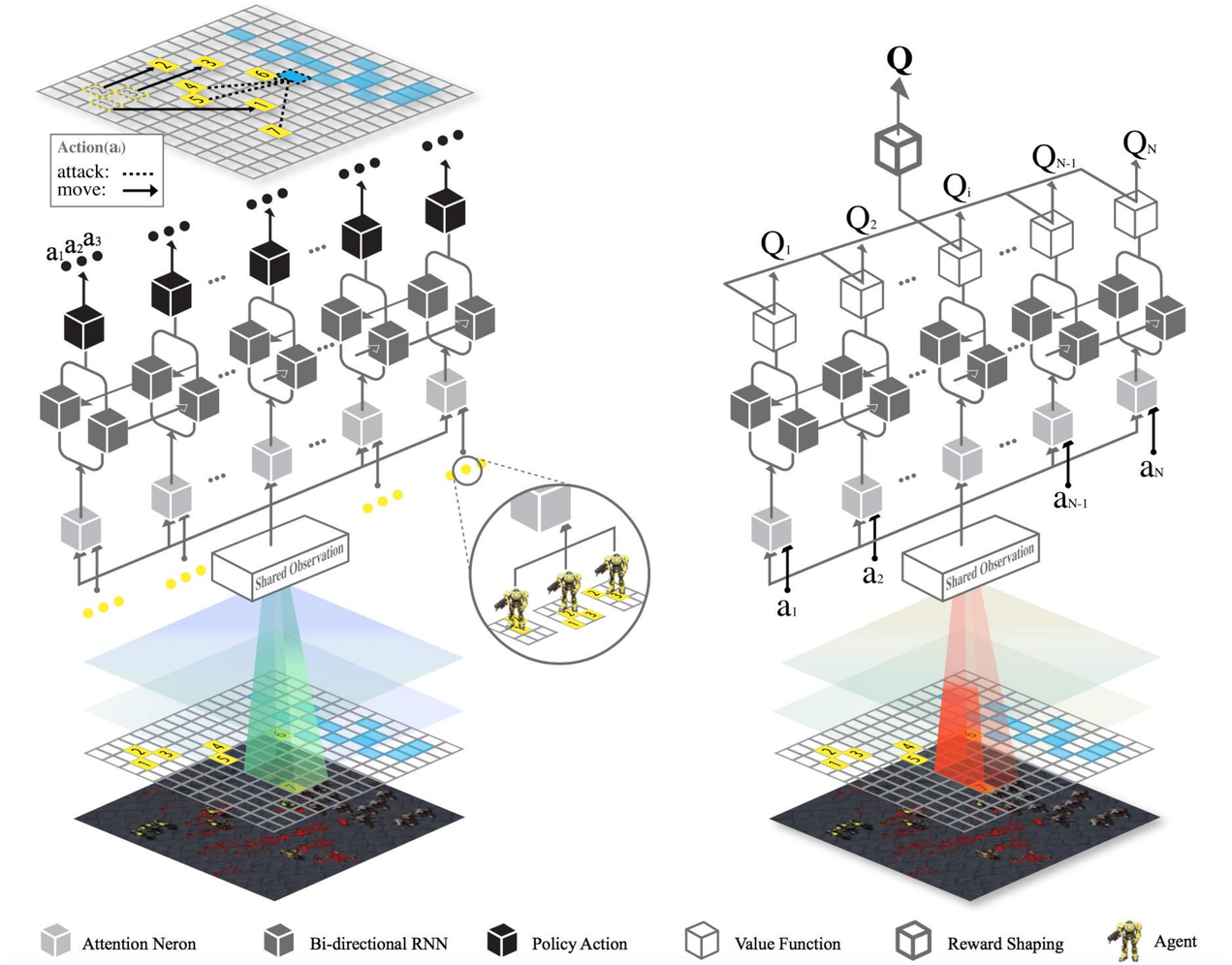
Peng Peng[†], **Quan Yuan**[†], **Ying Wen**[‡], **Yaodong Yang**[‡], **Zhenkun Tang**[†], **Haitao Long**[†], **Jun Wang**[‡] *
[†]Alibaba Group, [‡]University College London

Abstract

Real-world artificial intelligence (AI) applications often require multiple agents to work in a collaborative effort. Efficient learning for intra-agent communication and coordination is an indispensable step towards general AI. In this paper, we take StarCraft combat game as the test scenario, where the task is to coordinate multiple agents as a team to defeat their enemies. To maintain a scalable yet effective

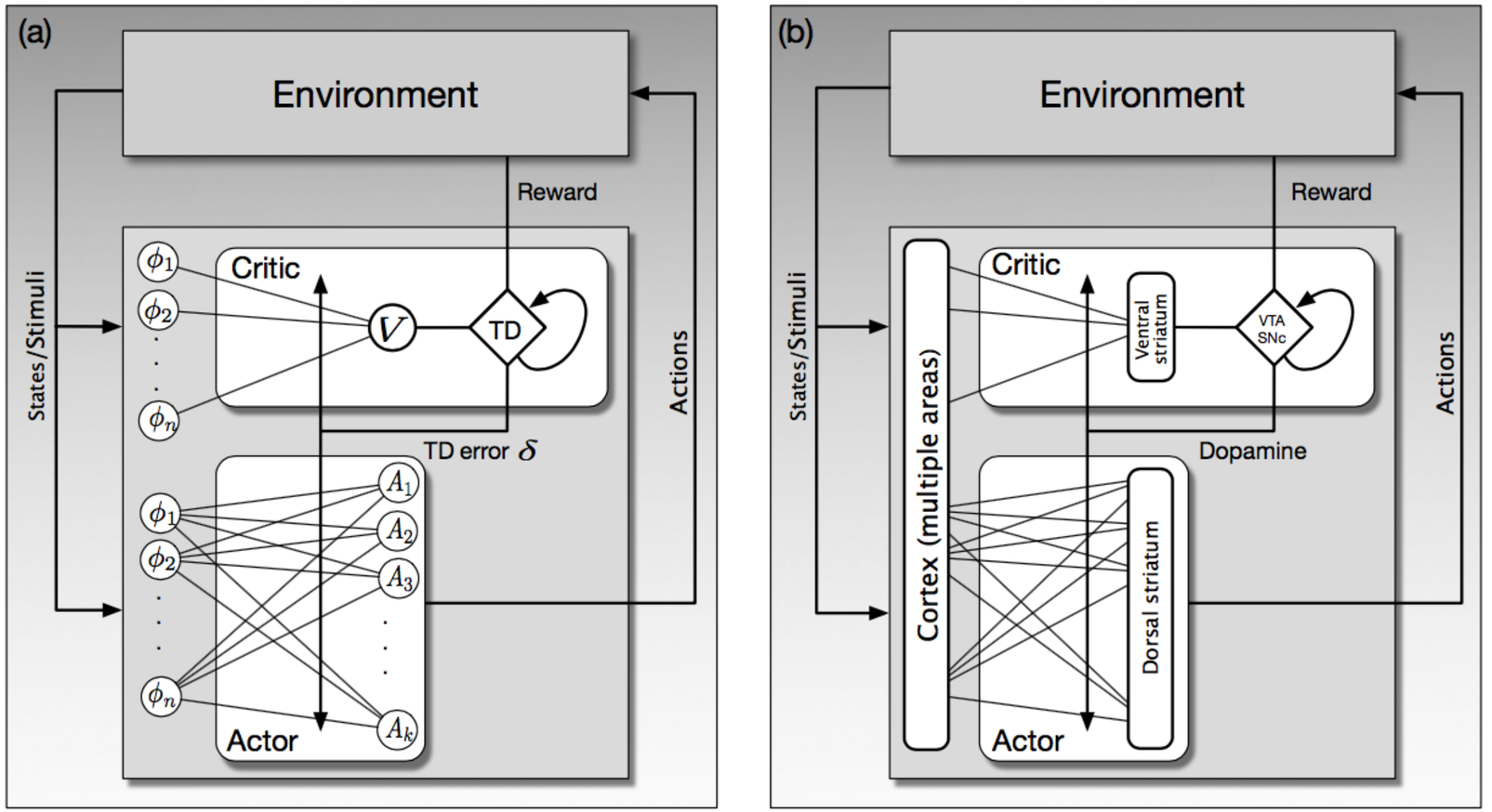
https://arxiv.org/abs/1703.10069

Multiagent Bidirectionally-Coordinated Net (BiCNet)



(a) Multiagent policy networks with grouping (b) Multiagent Q networks with reward shaping

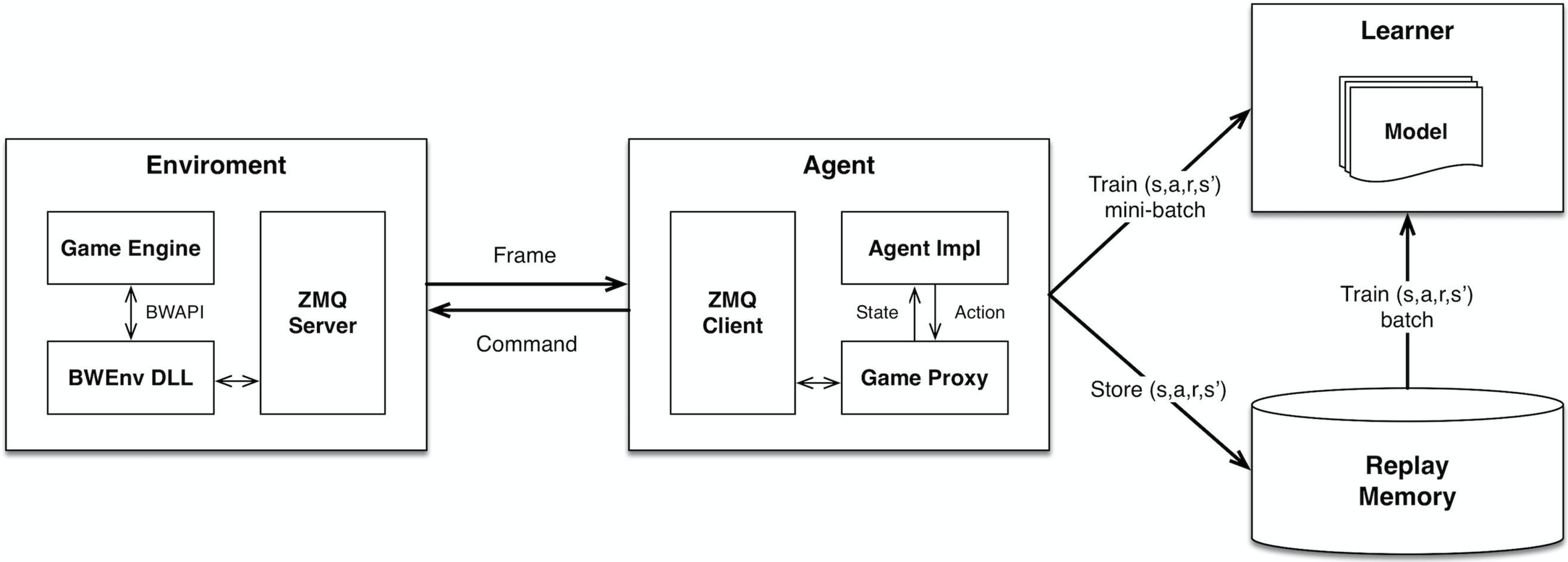
Neuroscience Hypothesis



Source: Reinforcement Learning - An Introduction

Alibaba Group 阿里巴巴集团

Architecture Overview



https://github.com/alibaba/gym-starcraft

Coordinated Moves without Collision

(a) Early stage of training

(b) Early stage of training

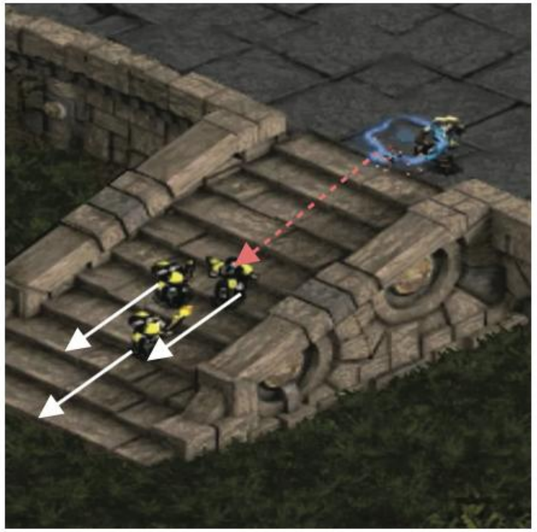
(c) Well-trained

(d) Well-trained

3 Marines (ours) vs. 1 Super Zergling (enemy)

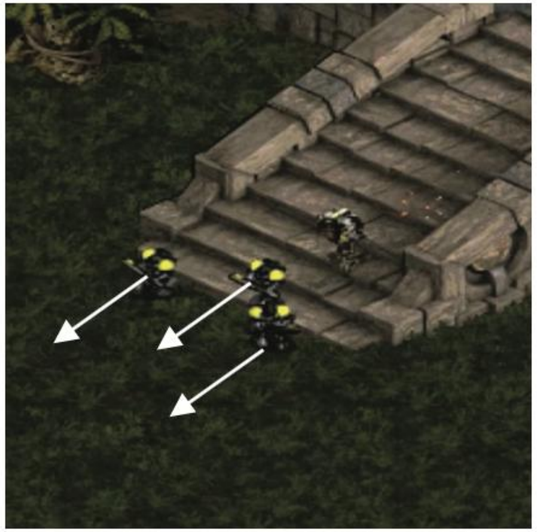
Alibaba Group
阿里巴巴集团

Hit and Run Tactics

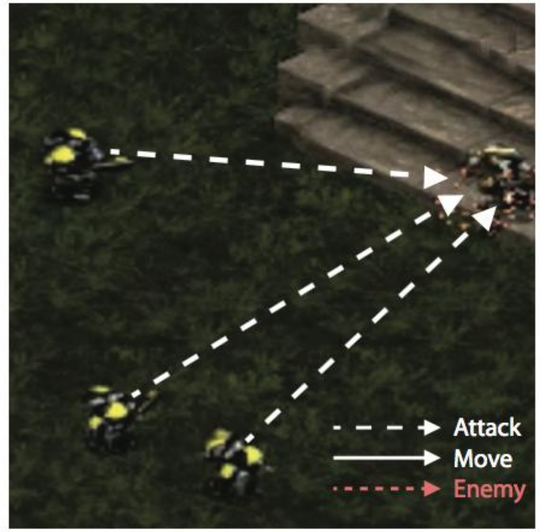


(a) time step 1: run when attacked

(b) time step 2: fight back when safe



(c) time step 3: run again

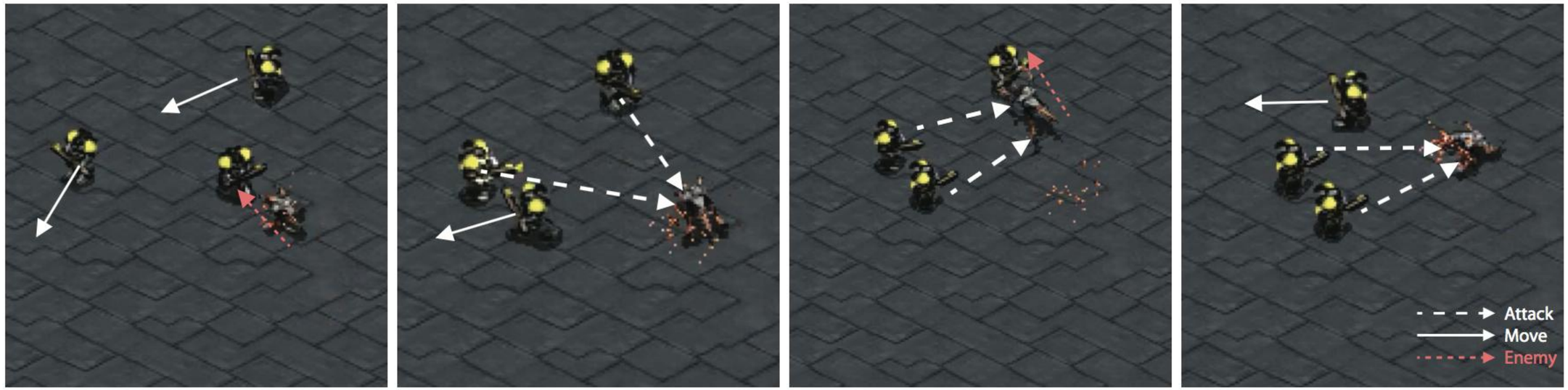


(d) time step 4: fight back again

3 Marines (ours) vs. 1 Zealot (enemy)

Alibaba Group
阿里巴巴集团

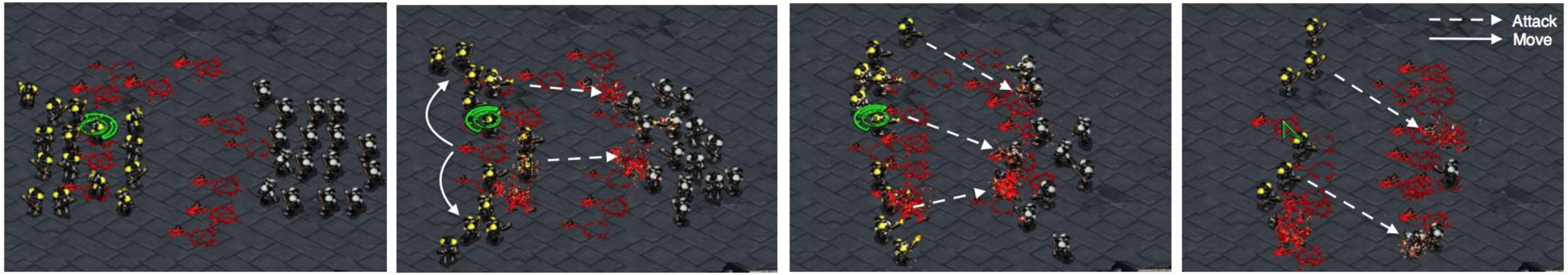
Coordinated Cover Attack



(a) time step 1 (b) time step 2 (c) time step 3 (d) time step 4

3 Marines (ours) vs. 1 Super Zergling (enemy)

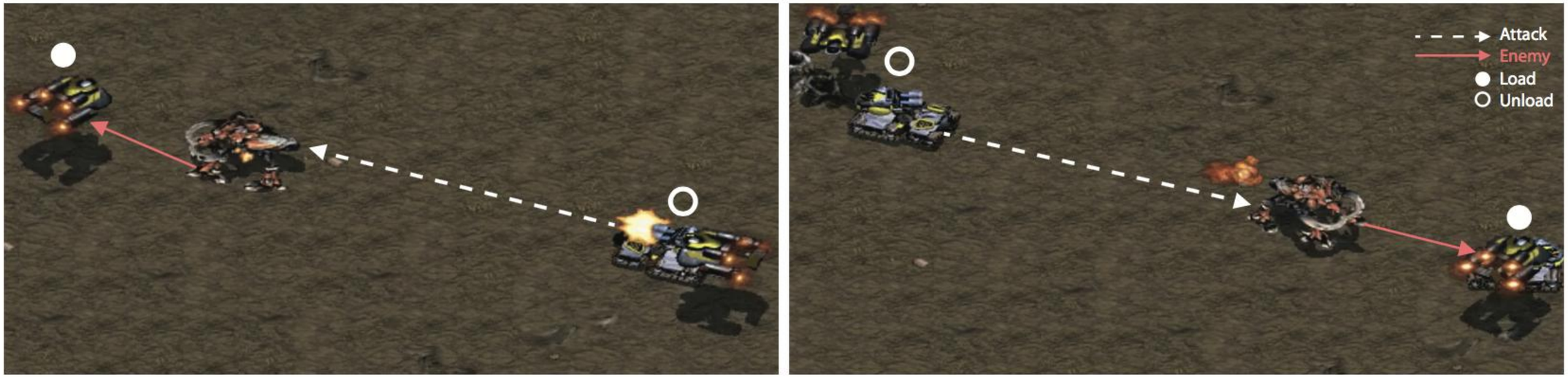
Focus Fire without Overkill



(a) time step 1 (b) time step 2 (c) time step 3 (d) time step 4

15 Marines (ours) vs. 16 Marines (enemy)

Coordinated Heterogonous Agents



(a) time step 1 (b) time step 2

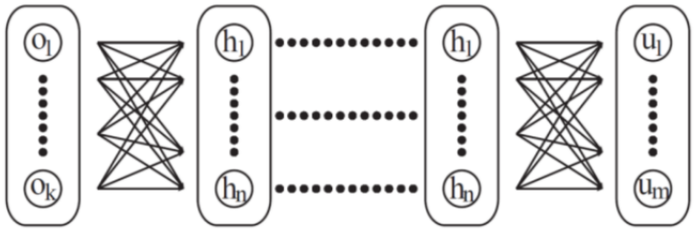
2 *Dropships* and 2 *tanks* vs. 1 *Ultralisk*

Alibaba Group 阿里巴巴集团

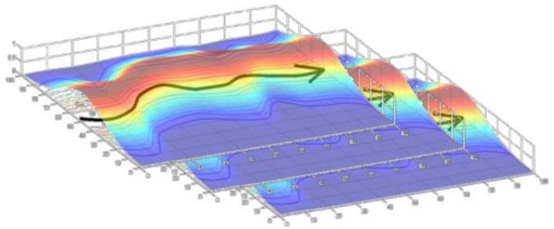
Hierarchical Reinforcement Learning

Alibaba Group
阿里巴巴集团

Imitation Learning

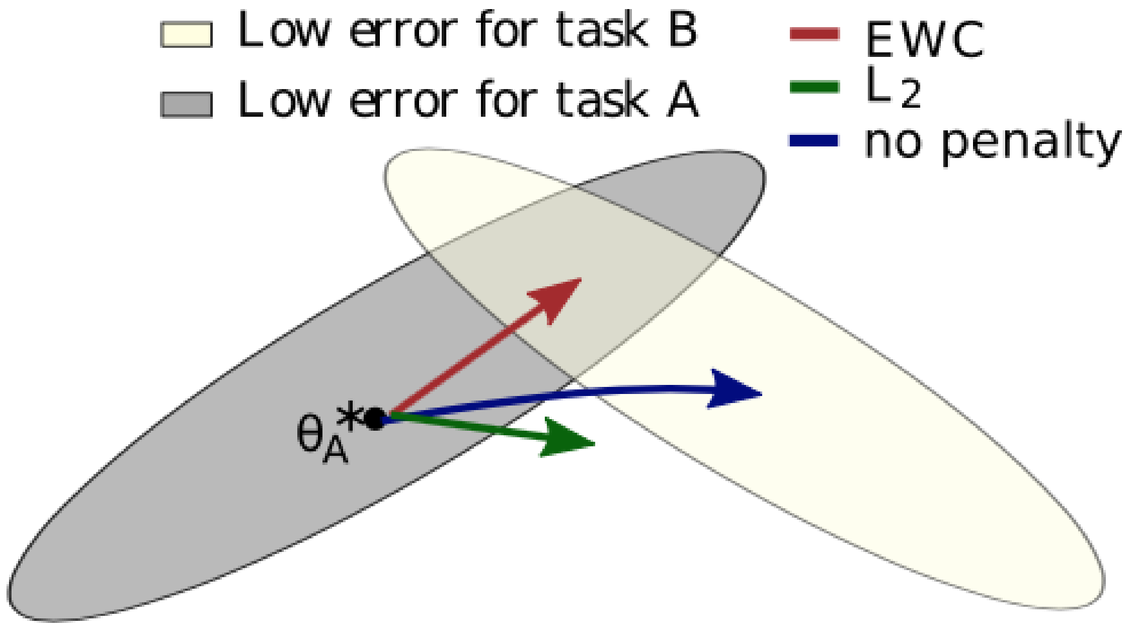


Supervised Learning



Reinforcement Learning

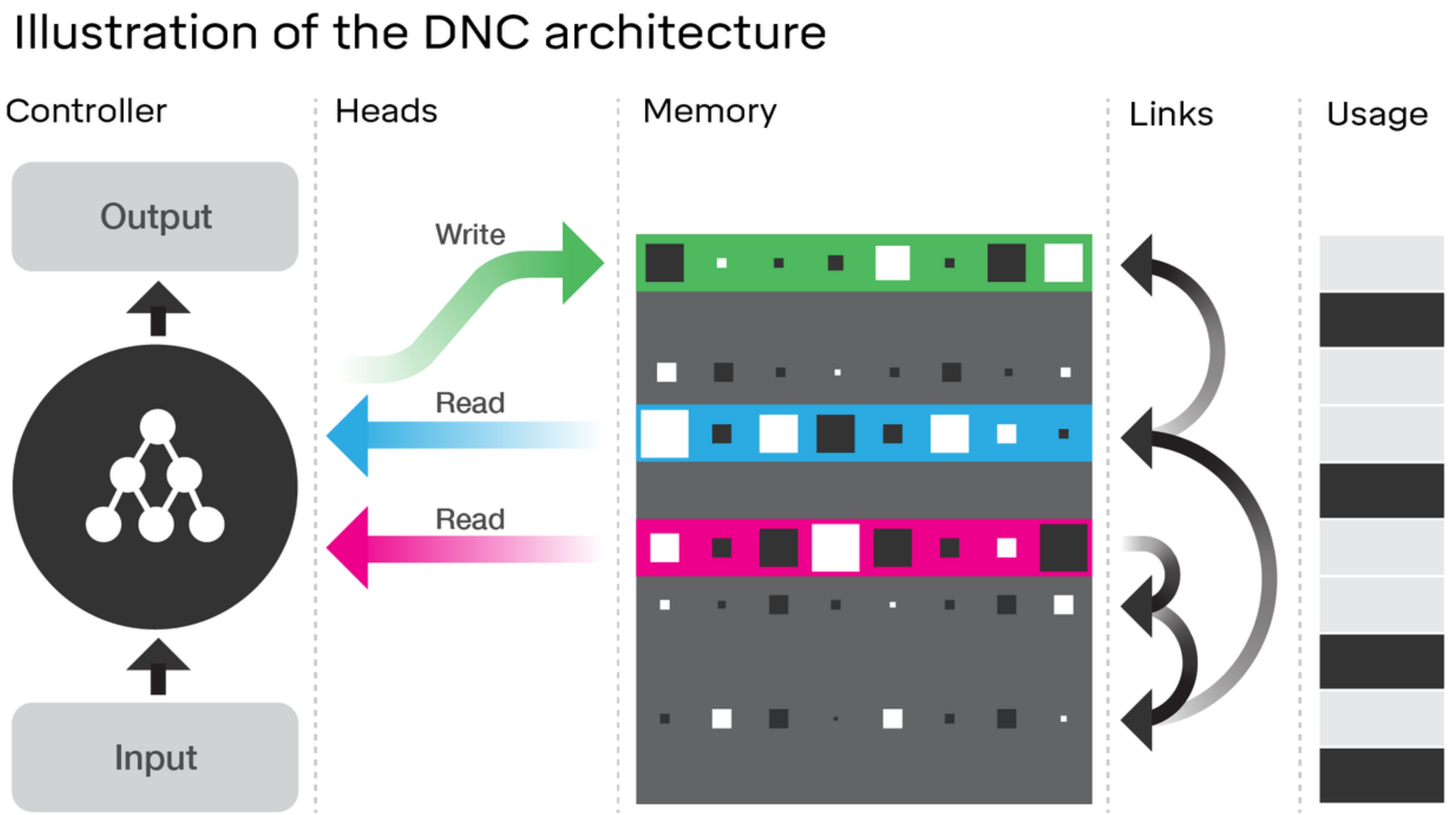
Continual Learning



Source: Overcoming catastrophic forgetting in neural networks

Alibaba Group
阿里巴巴集团

Memory-Augmented Neural Networks



Source: Hybrid computing using a neural network with dynamic external memory

Alibaba Group
阿里巴巴集团

Work Fun Play Hard