Hierarchical Reinforcement Learning

Introduction

Cliff Li

17.03.2020
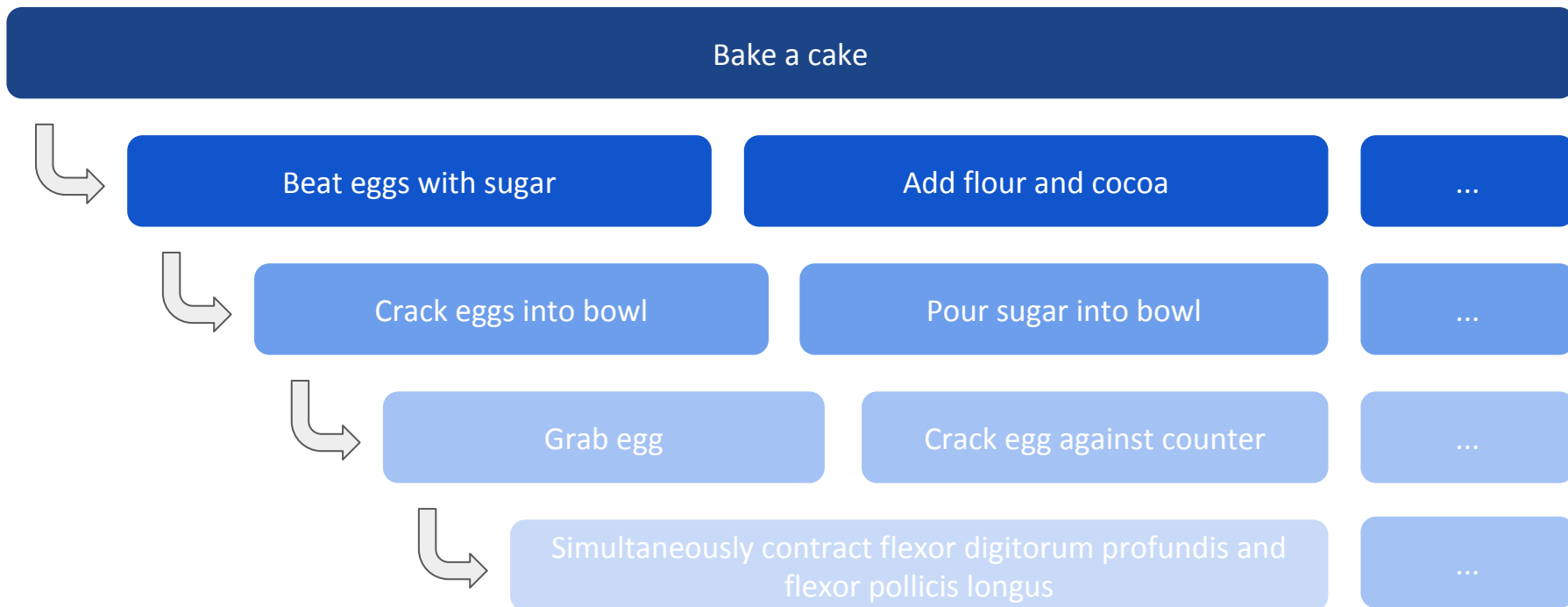
Cake

Source: reddit.com/u/GreekElisium

Deficiencies of classical Reinforcement Learning

- Huge state and action spaces

- Credit assignment

- Transfer learning

- Overfitting (Overspecialization)

- Knowledge representation

Hierarchical Reinforcement Learning in a nutshell

Bake a cake

Beat eggs with sugar | Add flour and cocoa | ...

Crack eggs into bowl | Pour sugar into bowl | ...

Grab egg | Crack egg against counter | ...

Simultaneously contract flexor digitorum profundis and flexor pollicis longus | ...
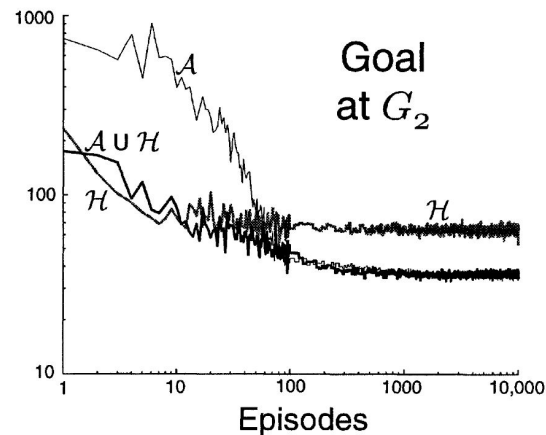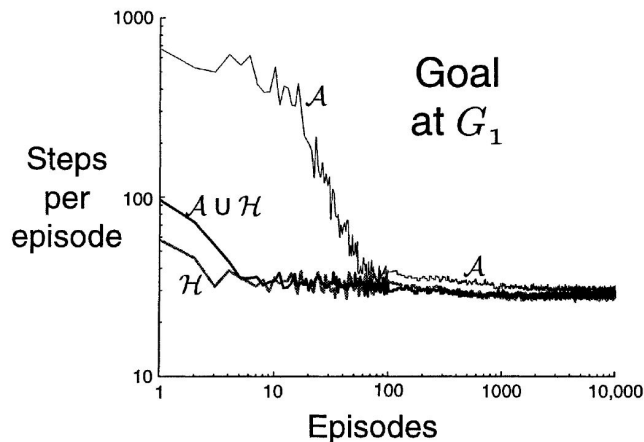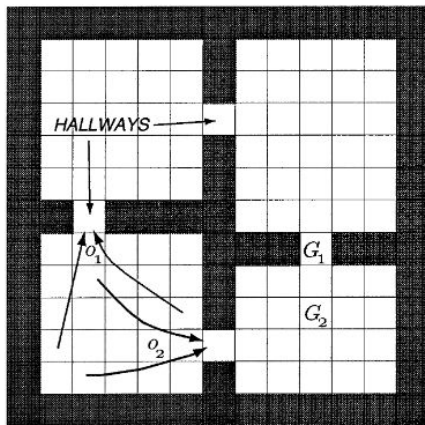
Benefits of Hierarchical Reinforcement Learning

- Structured exploration in state and action spaces

- Easier propagation of rewards

- Enables transfer learning

- Generalization through abstraction

- Better knowledge representation

Semi-Markov Decision Processes (SMDP) and Options
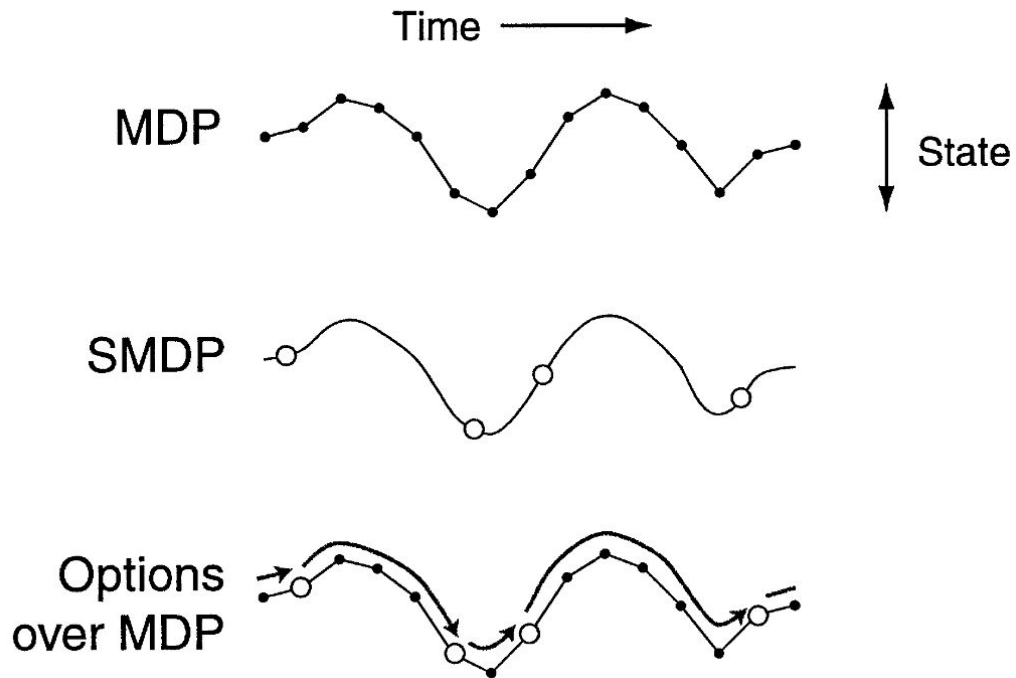
Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning (Sutton et al., 1999)

HALLWAYS

O_1

O_2

G_1

G_2

4 stochastic primitive actions

up

left ← → right   Fail 33% of the time

down

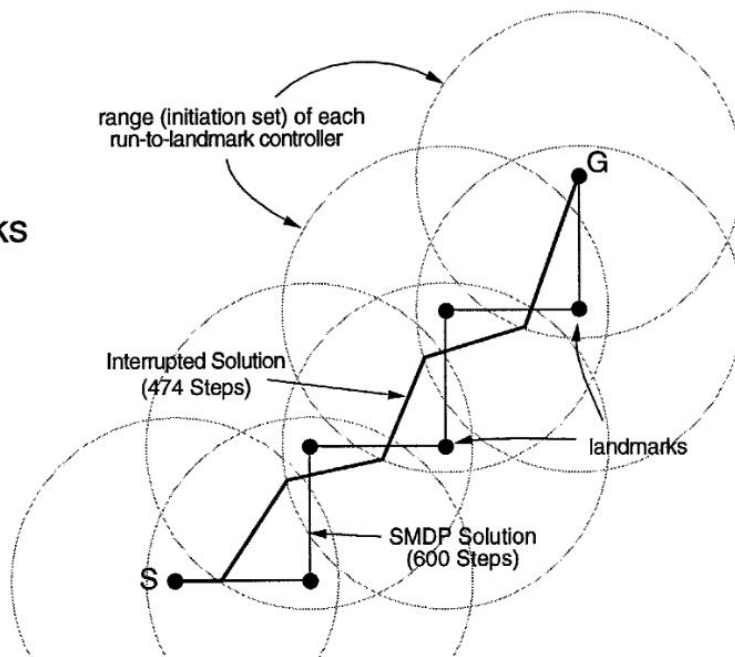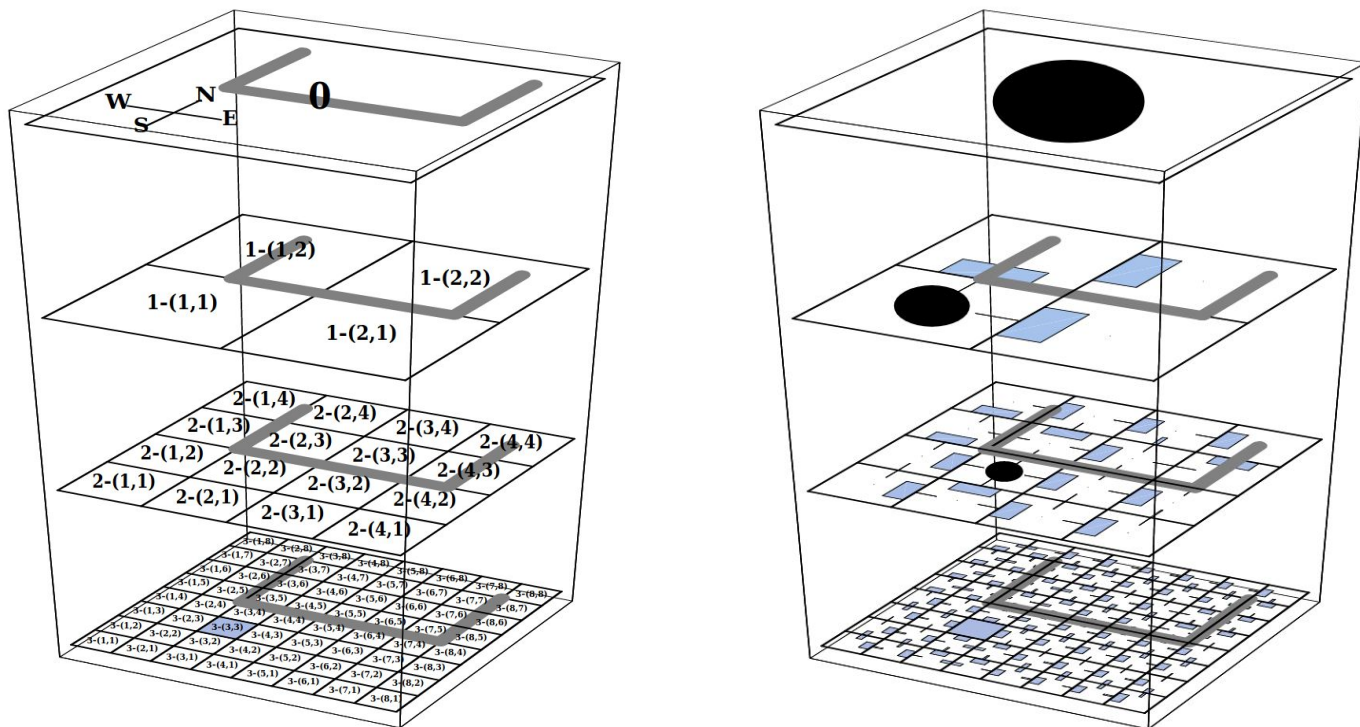8 multi-step options
(to each room's 2 hallways)

Target Hallway

Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning (Sutton et al., 1999)

Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning (Sutton et al., 1999)


Landmarks Problem

range (initiation set) of each run-to-landmark controller

Interrupted Solution (474 Steps)

landmarks

SMDP Solution (600 Steps)

G

S

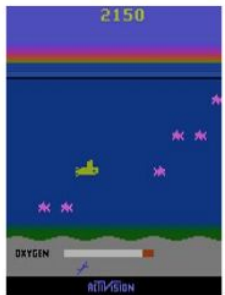# *Feudal Reinforcement Learning* (Dayan and Hinton, 1993)

Challenges of Hierarchical Reinforcement Learning

- Learning options

- Meaningful hierarchies

- Collapsing hierarchies into single policy

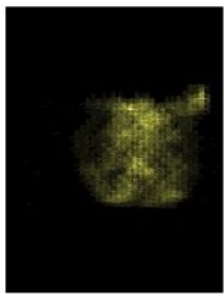- Updating lower-level policies affects higher-level performance

FeUdal Networks for Hierarchical Reinforcement Learning (Vezhnevets et al., 2017)
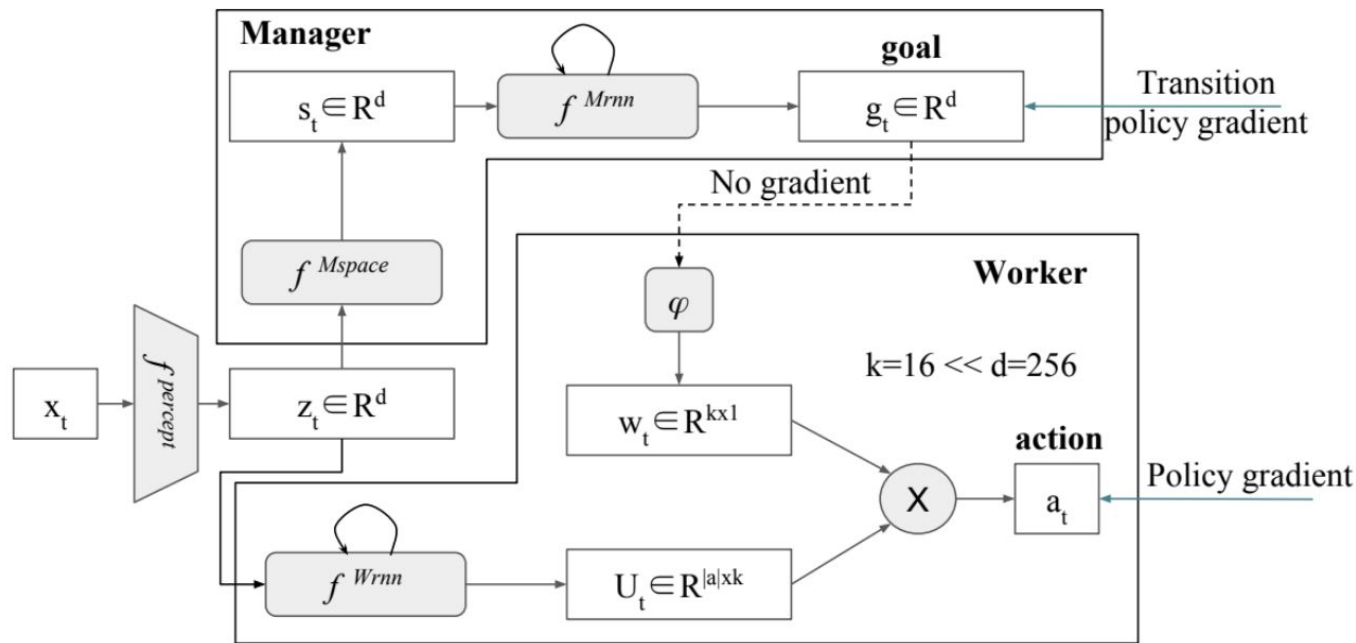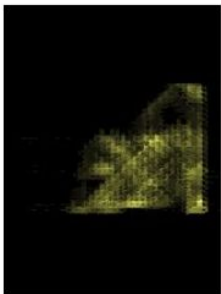
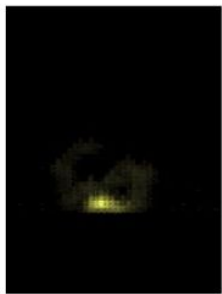# *FeUdal Networks for Hierarchical Reinforcement Learning* (Vezhnevets et al., 2017)
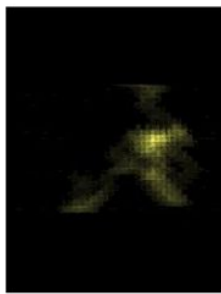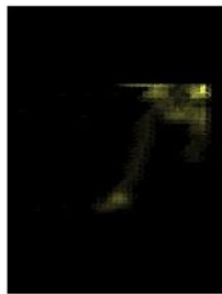

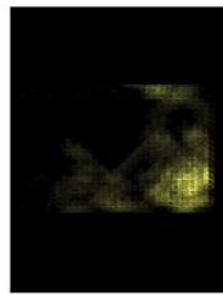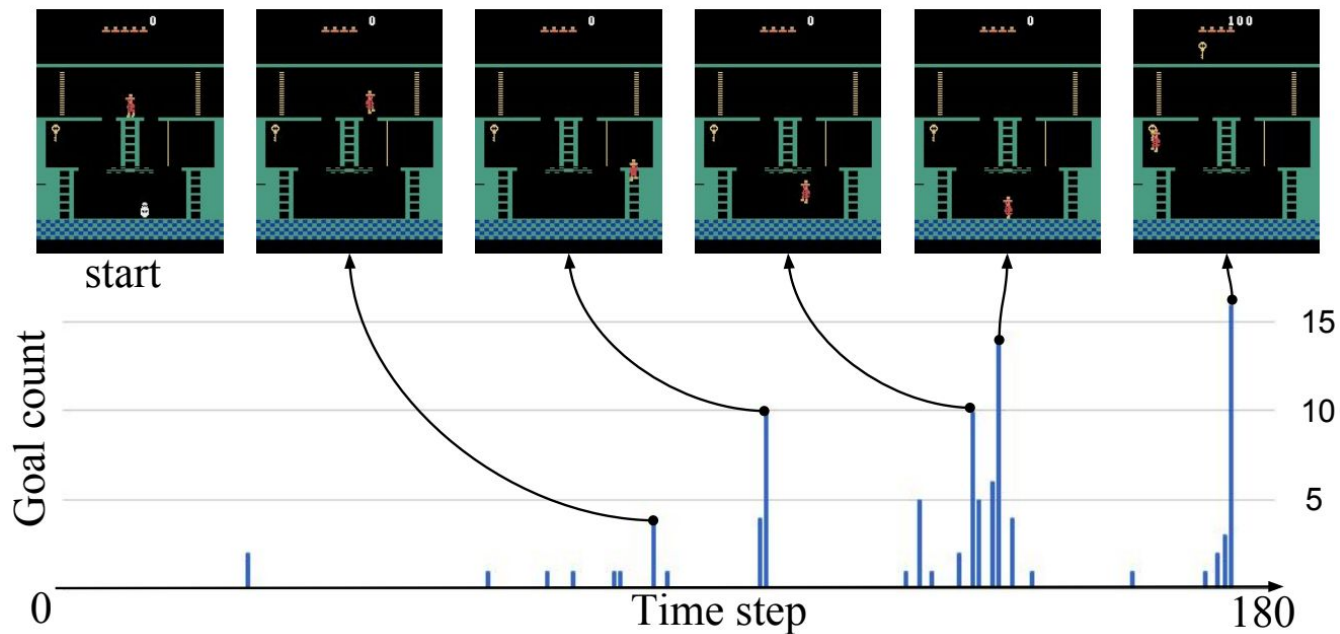Example frame   LSTM   Full FuN   sub-policy 1   sub-policy 2   sub-policy 3   sub-policy 4

FeUdal Networks for Hierarchical Reinforcement Learning (Vezhnevets et al., 2017)

References

Richard S. Sutton, Doina Precup, Satinder Singh. *Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning*. Artificial Intelligence, 1999.

Peter Dayan, Geoffrey Hinton. *Feudal Reinforcement Learning*. Advances in Neural Information Processing Systems, 1993.

Alexander Sasha Vezhnevets, Simon Osindero, Tom Schaul, Nicolas Heess, Max Jaderberg, David Silver, Koray Kavukcuoglu. *FeUdal Networks for Hierarchical Reinforcement Learning*. 2017.