

Parameterized Modular Inverse Reinforcement Learning

Shun Zhang

Chapter 1

Introduction

Chapter 2

Literature Review

2.1 Overview

Markov Decision Process MDP:

- State: S .
- Action: A .
- Transition: $P : S \times A \times S \rightarrow \mathcal{R}$.
- Reward: $R : S \times A \times S \rightarrow \mathcal{R}$.

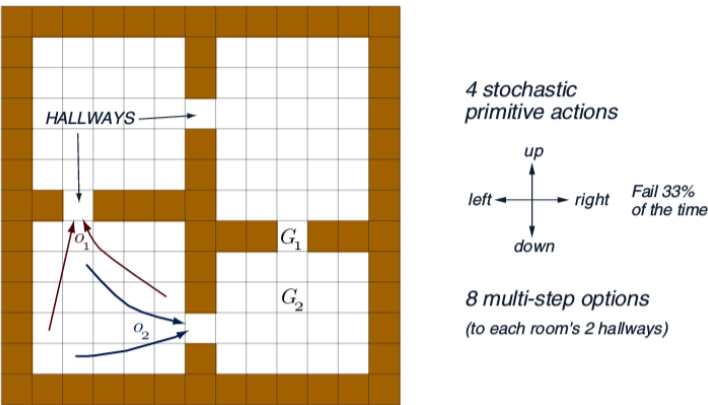
2.2 Forward Model

Abstraction on MDP

- Aggregate states: feature extraction.
- Aggregate actions: **option**.
- Decompose transition: factored MDP.
- Decompose value (abstract MDP): **HAM, hierarchical RL, modular RL**.

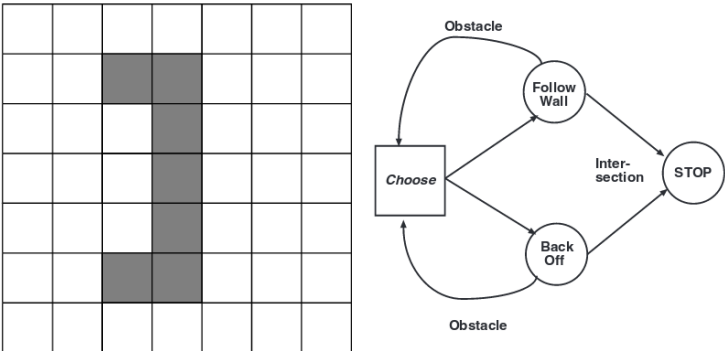
Approaches	State	Action	Reward	Value/Q
MDP with Option	/	Aggregated actions	/	/
Factored MDP	Decomposed states	/	/	Decomposed or not
HAM	/	/	/	/
Modular RL	/	/	/	/

Table 2.1: Overview of decomposition or aggregation of the components of MDP.



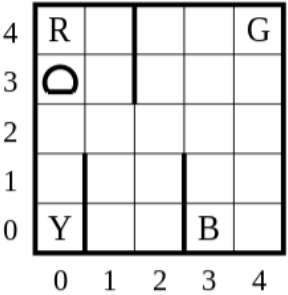
MDP with Option

- Option: (start state, policy, termination condition).
- State: S .
- Action: A, O .
- Transition: $P : S \times \{A, O\} \times S \rightarrow \mathcal{R}$.
- Reward: $R : S \times \{A, O\} \times S \rightarrow \mathcal{R}$.

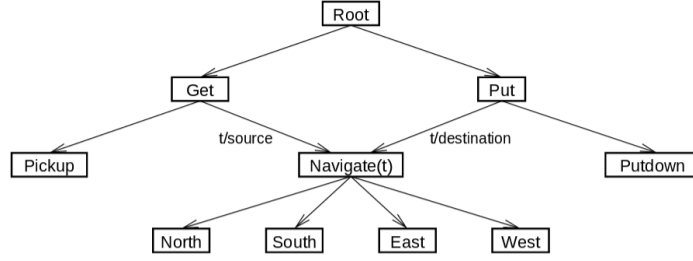


Hierarchies of Abstract Machines (HAM)

- State machine of MDPs.



Hierarchical RL



Hierarchical RL MDP:

- State: \mathcal{S} .
- Action: \mathcal{A} .
- Transition: \mathcal{T} .
- Reward: \mathcal{R} .

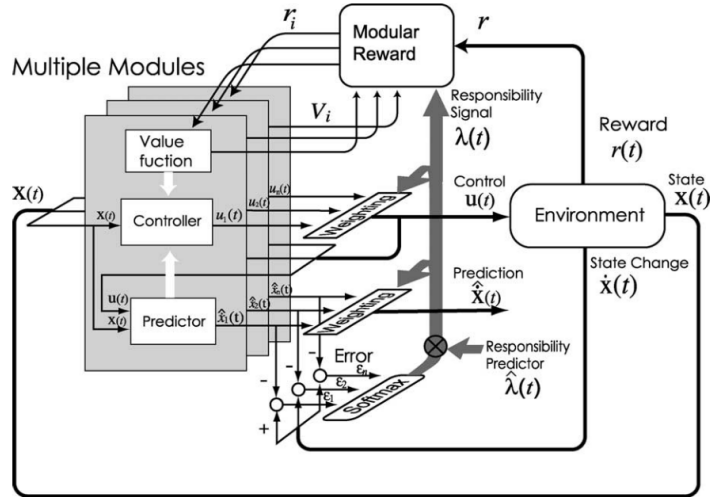


Fig. 1. MMRL.

Modular RL
MDP:

- State: $\mathcal{S}_1 \times \mathcal{S}_2 \cdots \times \mathcal{S}_M$.
- Action: \mathcal{A} .
- Transition: $\mathcal{P}_1 \times \mathcal{P}_2 \cdots \times \mathcal{P}_M$.
- Reward: $\mathcal{R}_1 \times \mathcal{R}_2 \cdots \times \mathcal{R}_M$.

2.3 Ideas in Recent Work

Not fix a model.

Learn the components,

Dynamic.

Chapter 3

Parameterized Modular Inverse Reinforcement Learning

3.1 Modular Inverse Reinforcement Learning

Chapter 4

Application: Human Behavior Modeling

4.1 Experiments

4.1.1 Grid World

4.1.2 Human Motion

Chapter 5

Application: Influencing Agents in a Flock

5.1 Introduction

Weighted sum of actions [1].

Chapter 6

Conclusion

Bibliography

- [1] Katie Genter, Shun Zhang, and Peter Stone. Determining placements of influencing agents in a flock. 2015.