

Survey on Hierarchical and Modular Reinforcement Learning

Shun Zhang

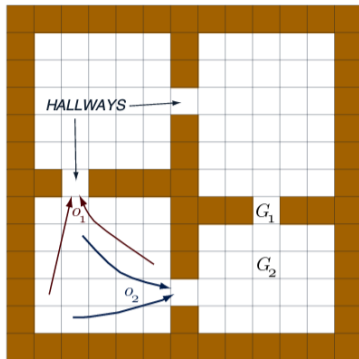
September 25, 2014

MDP:

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

- Aggregate states: feature extraction.
- Aggregate actions: **option**.
- Factorize transition: factored MDP.
- Abstract MDP: **HAM, hierarchical RL, modular RL**.

MDP with Option



*4 stochastic
primitive actions*

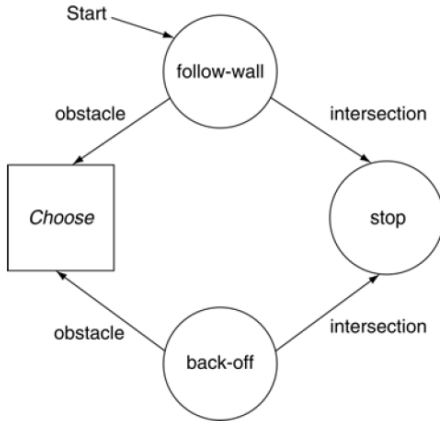


*8 multi-step options
(to each room's 2 hallways)*

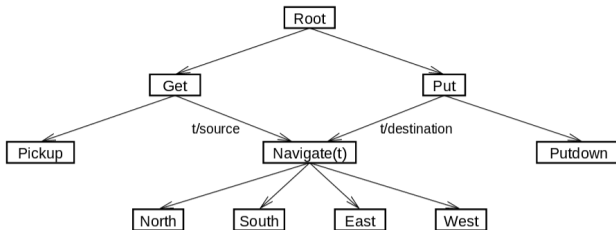
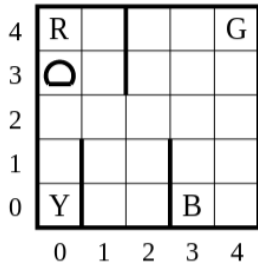
MDP:

- 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)



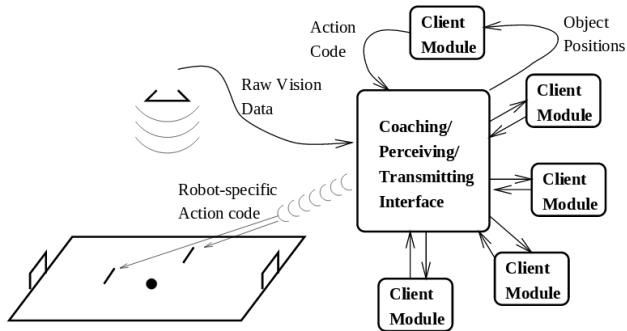
Hierarchical RL



MDP:

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

Layered RL



MDP:

- State: $S_1 \times S_2 \cdots \times S_M$.
- Action: $A_1 \times A_2 \cdots \times A_M$.
- Transition: $P_1 \times P_2 \cdots \times P_M$.
- Reward: $R_1 \times R_2 \cdots \times R_M$.

Modular RL

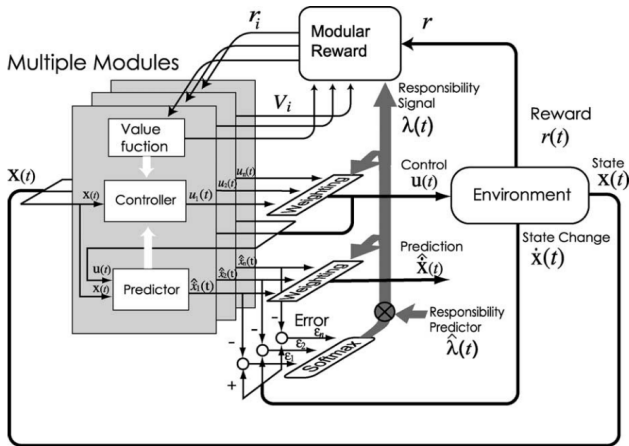


Fig. 1. MMRL.

MDP:

- State: $S_1 \times S_2 \cdots \times S_M$.
- Action: A .
- Transition: $P_1 \times P_2 \cdots \times P_M$.
- Reward: $R_1 \times R_2 \cdots \times R_M$.

- Learning task hierarchies.
- Dynamic abstraction.
- Integrating Deep Learning.