Opponent Modeling based on Subgoal Inference

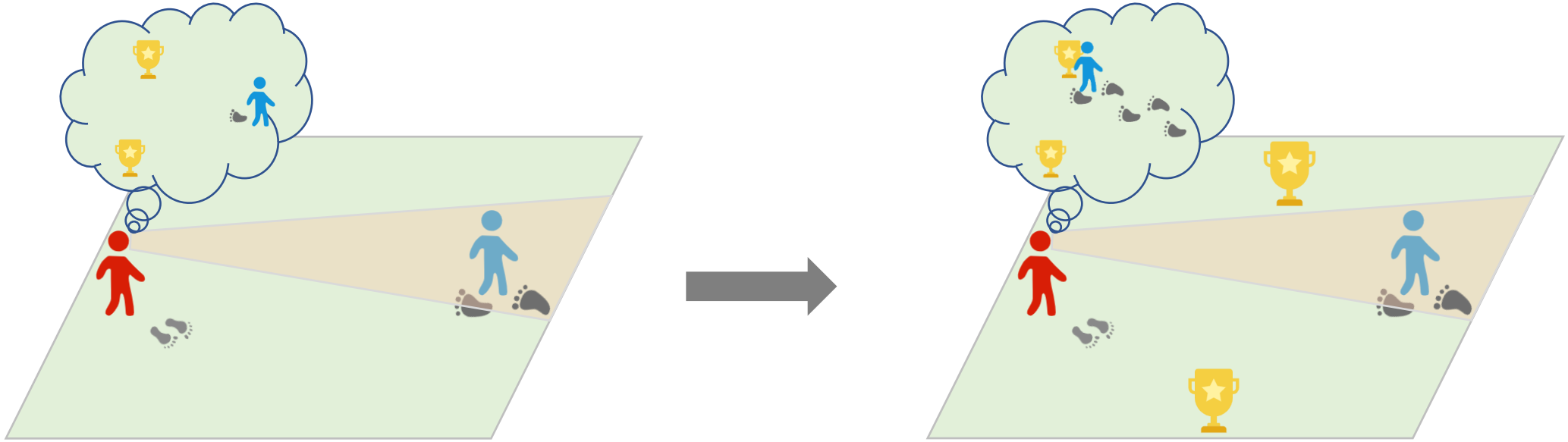
Xiaopeng Yu, Jiechuan Jiang, Zongqing Lu

School of Computer Science, Peking University

NeurIPS 2024

Introduction

Opponent **M**odeling based on sub**G**oal inference (**OMG**)

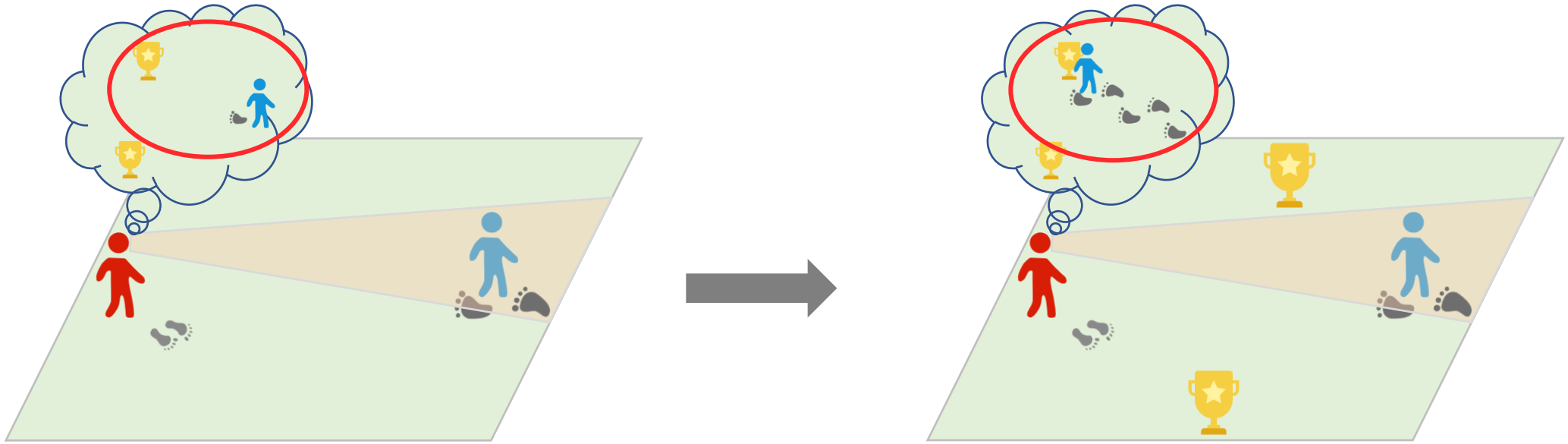


Modeling opponent's action

Modeling opponent's subgoal

Introduction

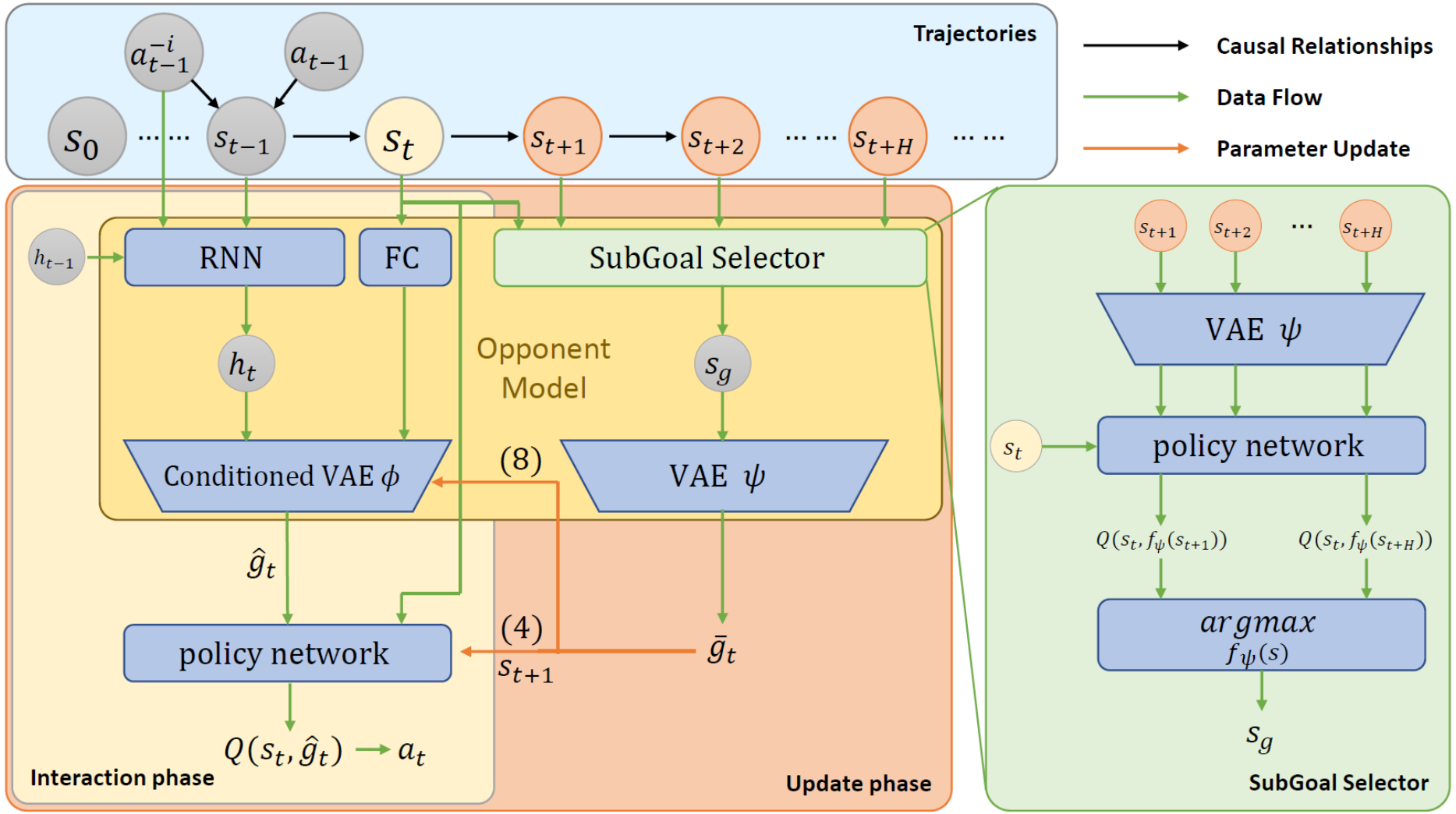
Opponent **M**odeling based on sub**G**oal inference (**OMG**)



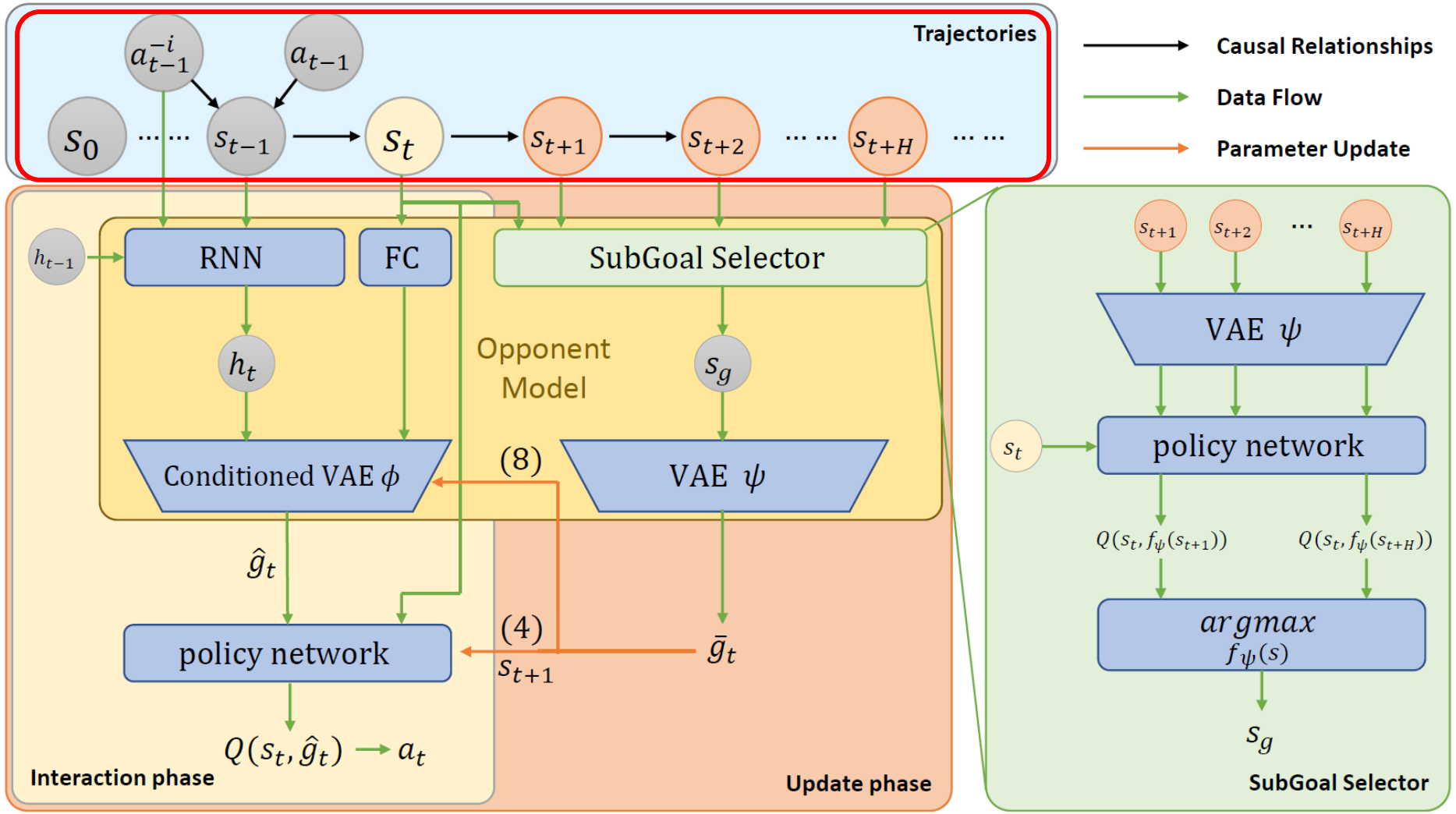
Modeling opponent's action

Modeling opponent's subgoal

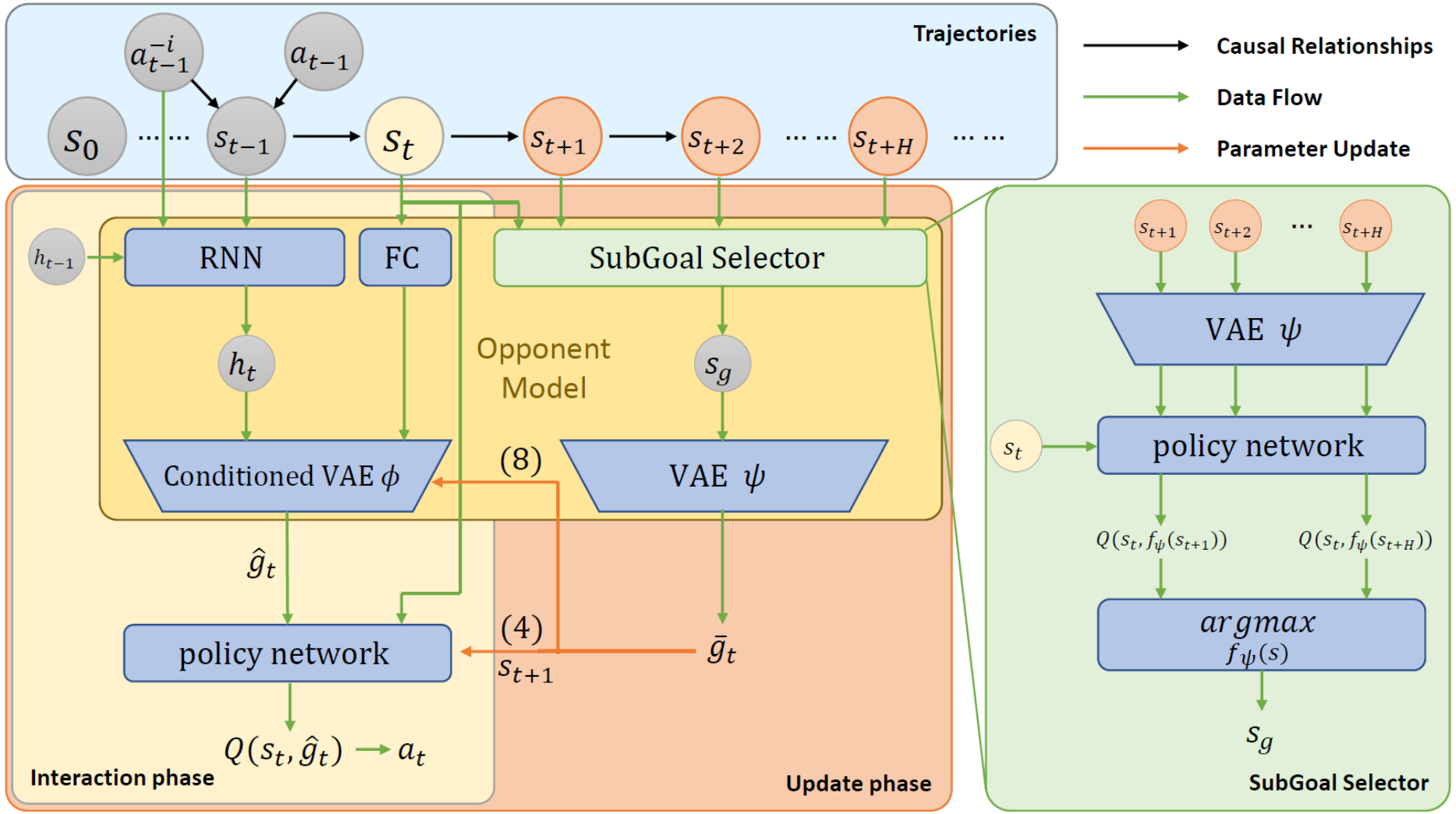
Method


Trajectories
a_{t-1}^{-i}
a_{t-1}
s_0
s_{t-1}
s_t
s_{t+1}
s_{t+2}
s_{t+H}
Causal Relationships
Data Flow
Parameter Update
h_{t-1}
RNN
FC
SubGoal Selector
h_t
Opponent Model
s_g
Conditioned VAE ϕ
(8)
VAE ψ
$\hat{g}_t$
(4)
$\bar{g}_t$
policy network
s_{t+1}
$Q(s_t, \hat{g}_t) \rightarrow a_t$
Interaction phase
Update phase
VAE ψ
s_t
policy network
$Q(s_t, f_\psi(s_{t+1}))$
$Q(s_t, f_\psi(s_{t+H}))$
$argmax$ $f_\psi(s)$
s_g
SubGoal Selector

Method

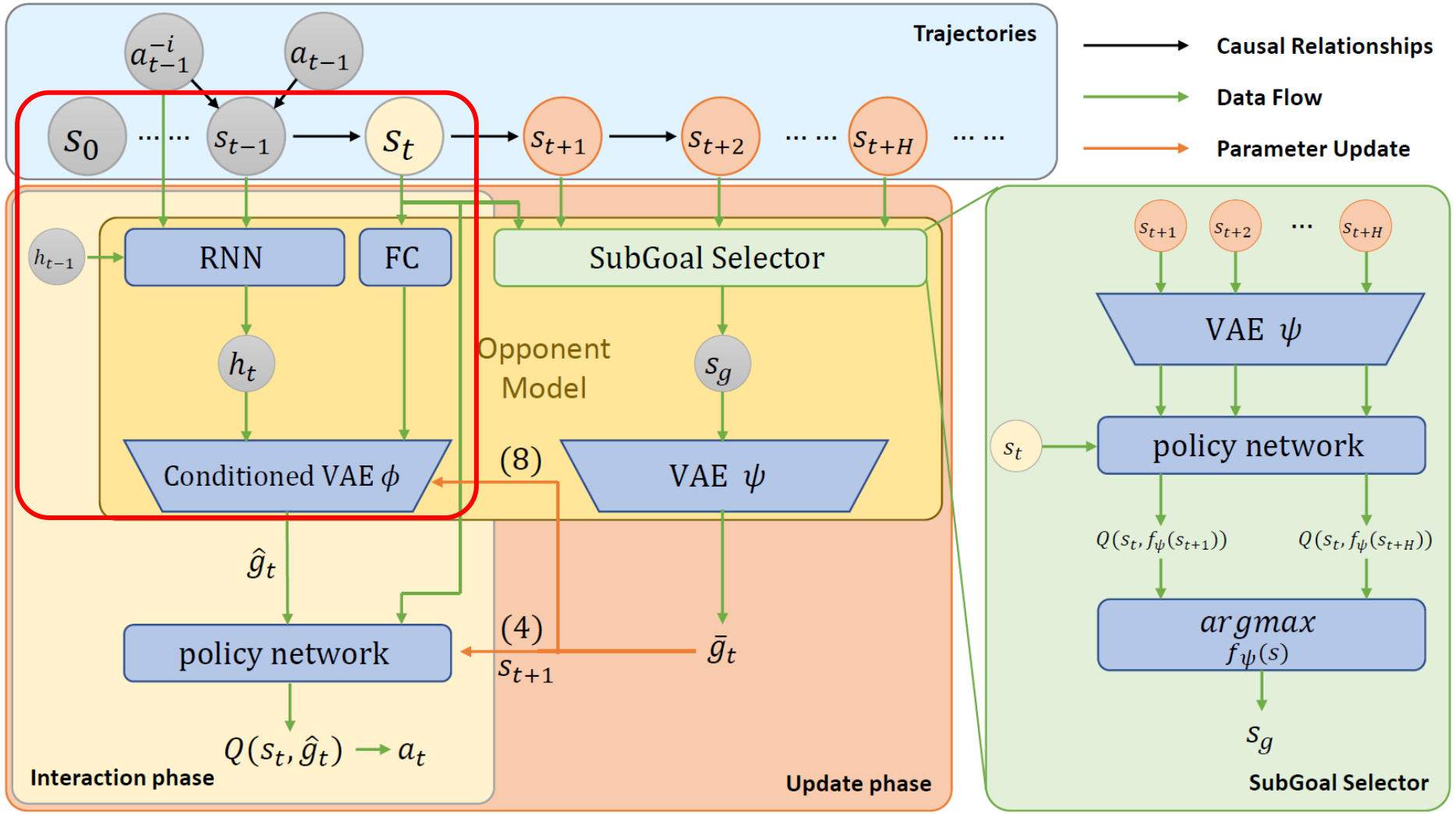

Trajectories
a_{t-1}^{-i}
a_{t-1}
s_0
s_{t-1}
s_t
s_{t+1}
s_{t+2}
s_{t+H}
Causal Relationships
Data Flow
Parameter Update
h_{t-1}
RNN
FC
SubGoal Selector
h_t
Opponent Model
s_g
Conditioned VAE ϕ
(8)
VAE ψ
$\hat{g}_t$
(4)
$\bar{g}_t$
policy network
s_{t+1}
$Q(s_t, \hat{g}_t) \rightarrow a_t$
Interaction phase
Update phase
VAE ψ
s_t
policy network
$Q(s_t, f_\psi(s_{t+1}))$
$Q(s_t, f_\psi(s_{t+H}))$
$argmax$ $f_\psi(s)$
s_g
SubGoal Selector

Method


Trajectories
a_{t-1}^{-i}
a_{t-1}
s_0
s_{t-1}
s_t
s_{t+1}
s_{t+2}
s_{t+H}
Causal Relationships
Data Flow
Parameter Update
h_{t-1}
RNN
FC
SubGoal Selector
h_t
Opponent Model
s_g
Conditioned VAE ϕ
(8)
VAE ψ
$\hat{g}_t$
(4)
$\bar{g}_t$
policy network
s_{t+1}
$Q(s_t, \hat{g}_t) \rightarrow a_t$
Interaction phase
Update phase
VAE ψ
s_t
policy network
$Q(s_t, f_\psi(s_{t+1}))$
$Q(s_t, f_\psi(s_{t+H}))$
$argmax$ $f_\psi(s)$
s_g
SubGoal Selector

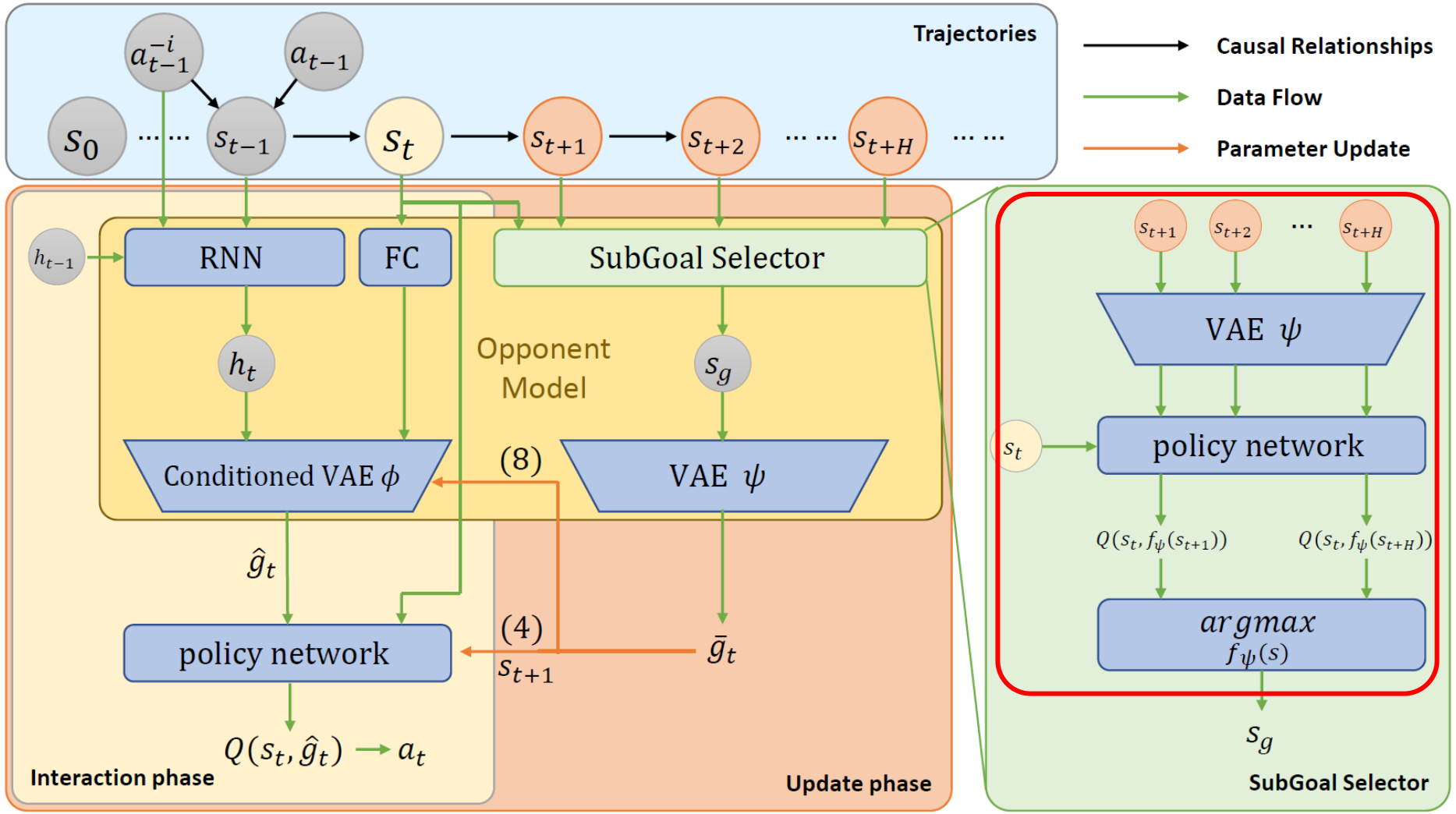
Method

Subgoal Inference Model



Method

Subgoal Selector



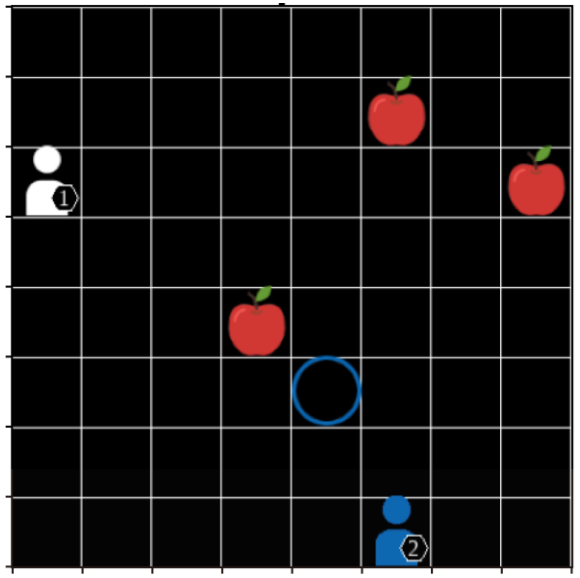
$$\bar{g}_t = \arg\max_{s_i \in \mathcal{N}_t^H} \mathbb{E}_{g \sim p_\psi(\cdot|s_i)} V(s_t, g)$$

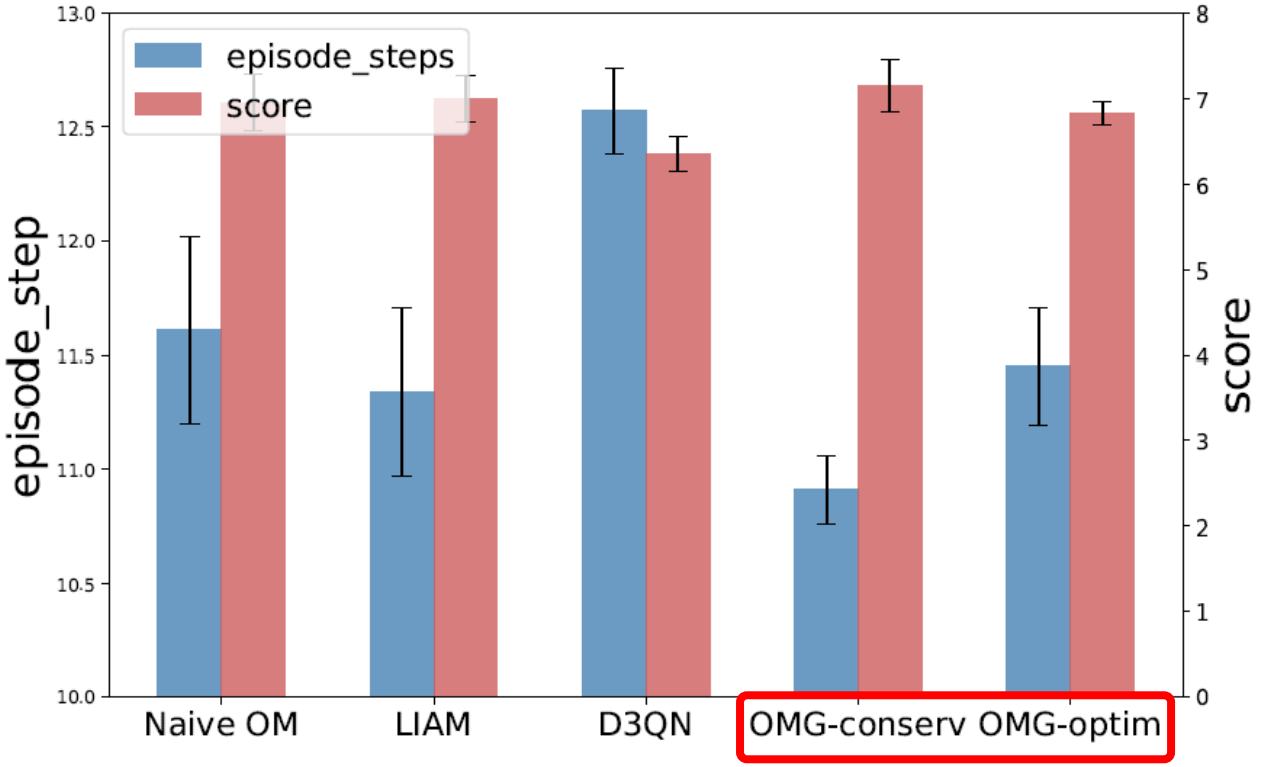
$$\bar{g}_t = \arg\min_{s_i \in \mathcal{N}_t^H} \mathbb{E}_{g \sim p_\psi(\cdot|s_i)} V(s_t, g),$$

Experiment

Competitive Scenarios

Foraging

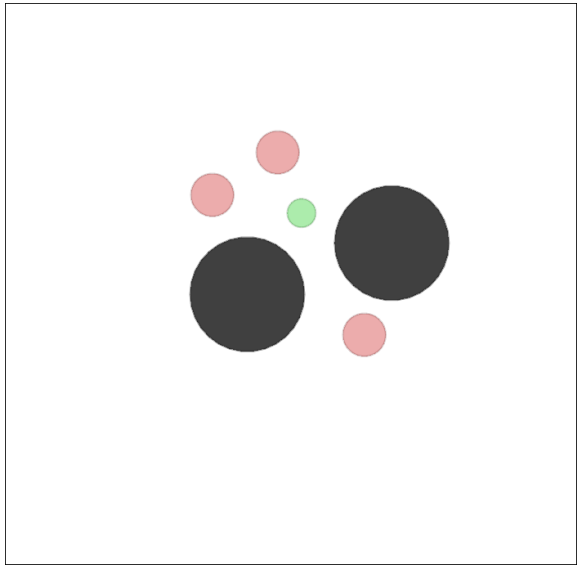


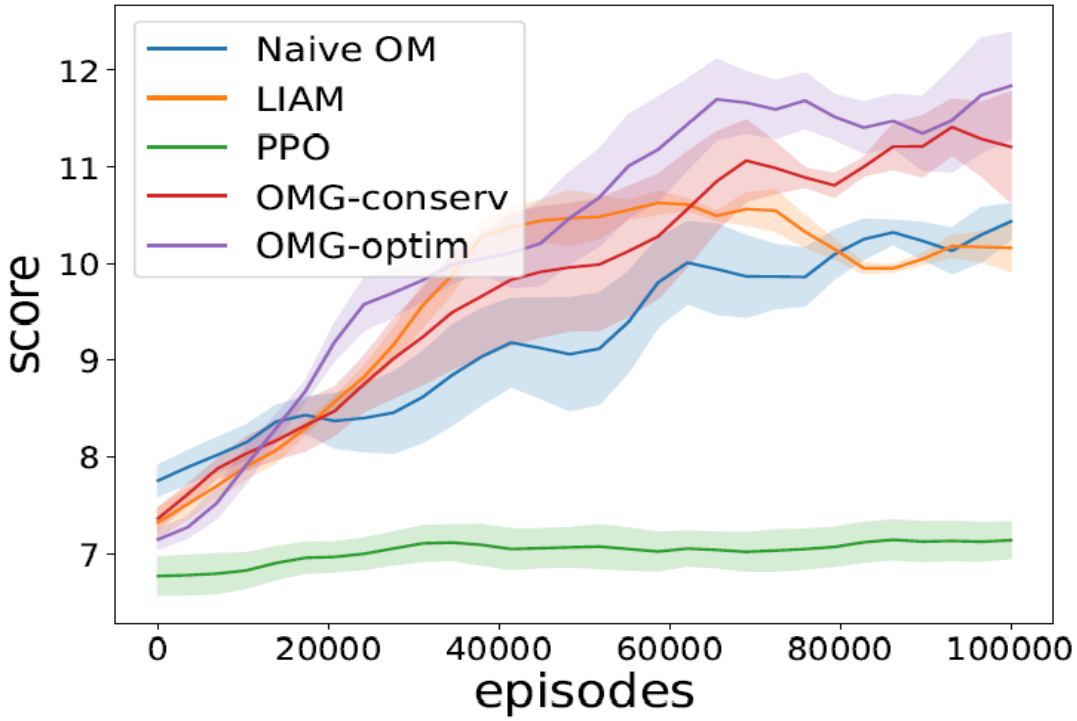


Experiment

Competitive Scenarios

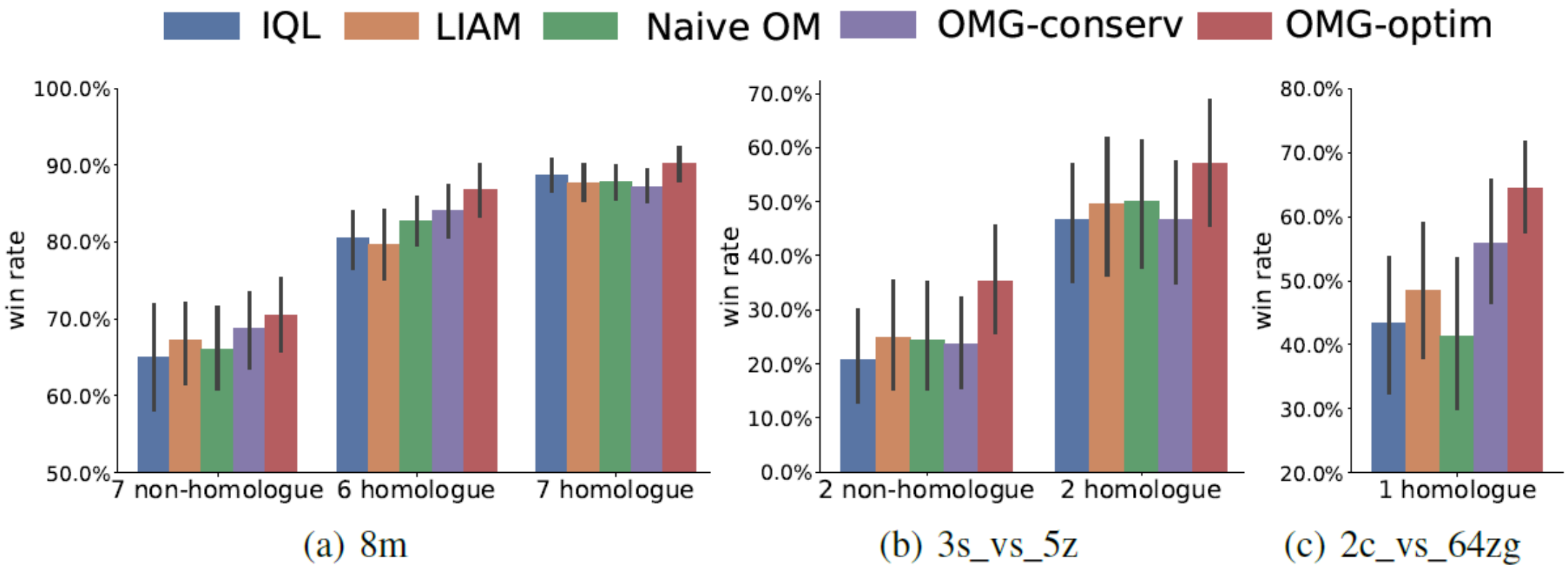
Predator-Prey





Experiment

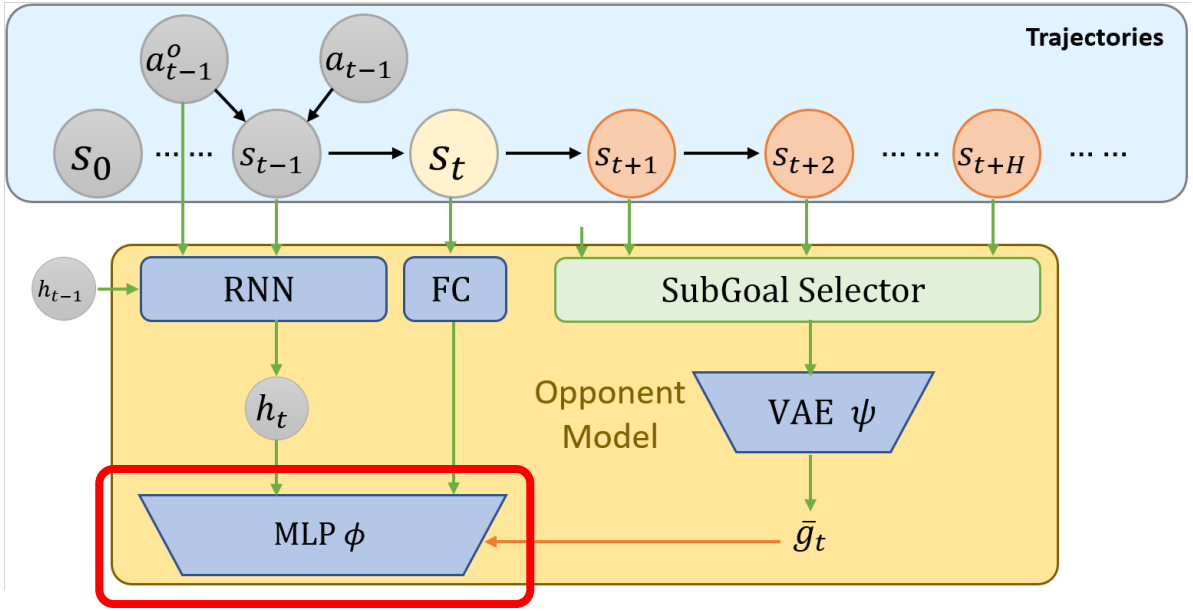
Cooperative Settings

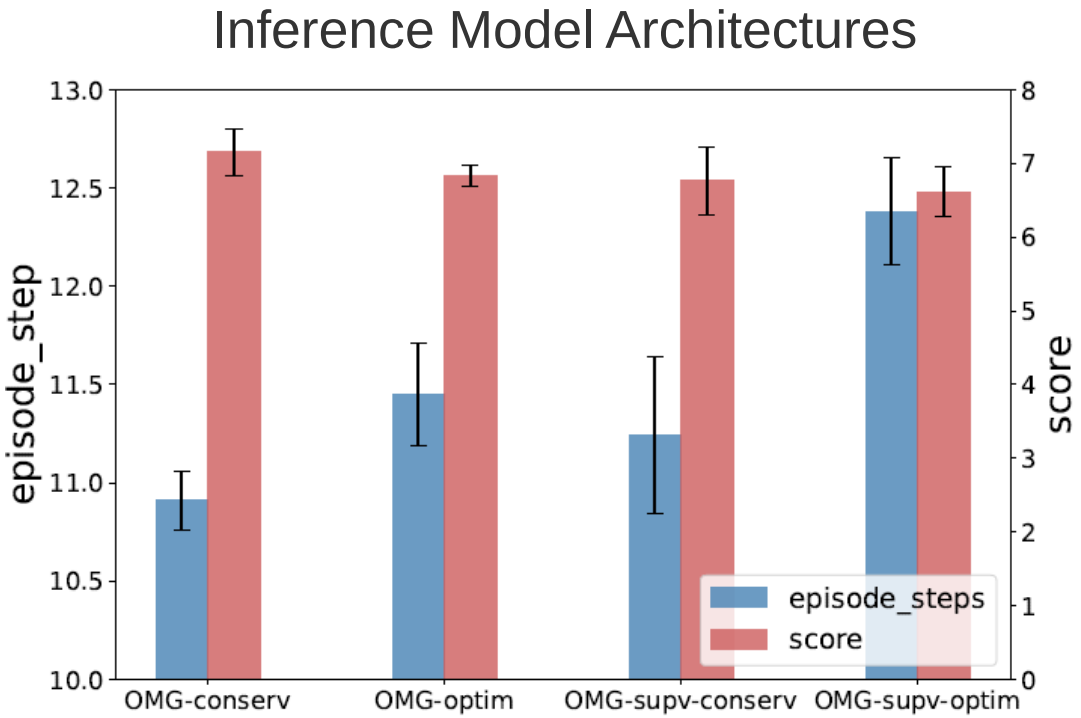


SMAC

Experiment

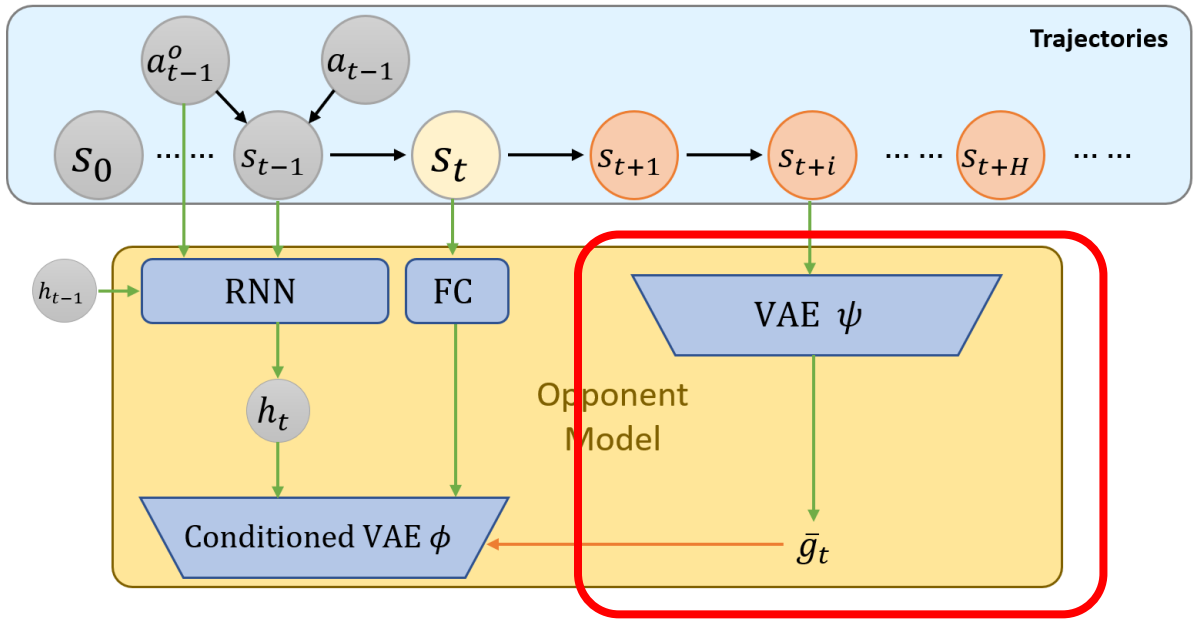
Ablation Study



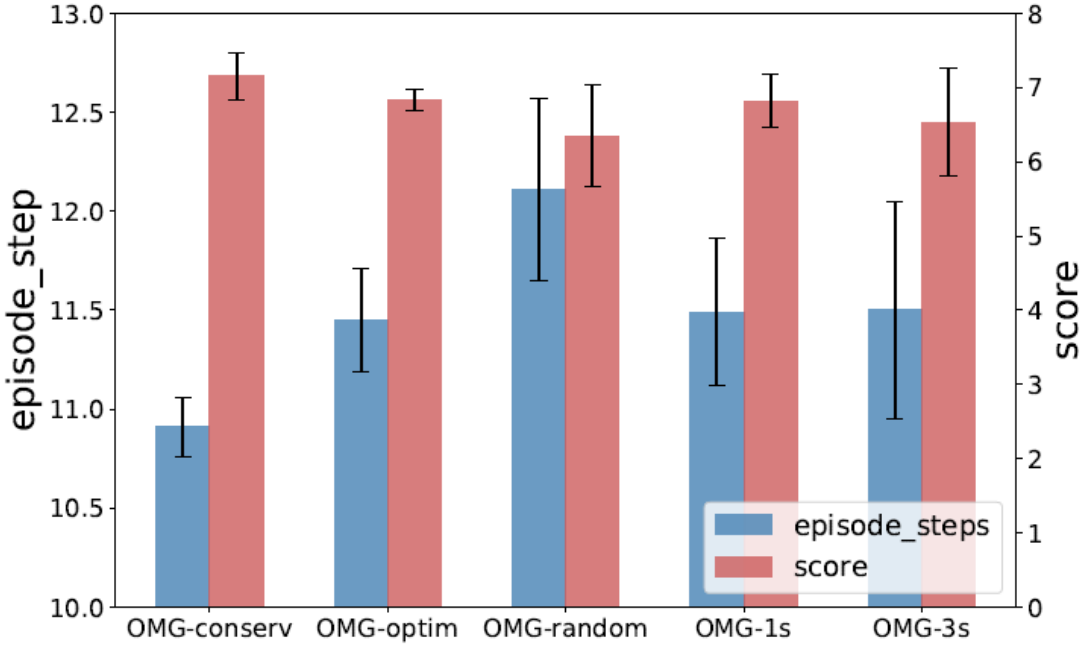


Experiment

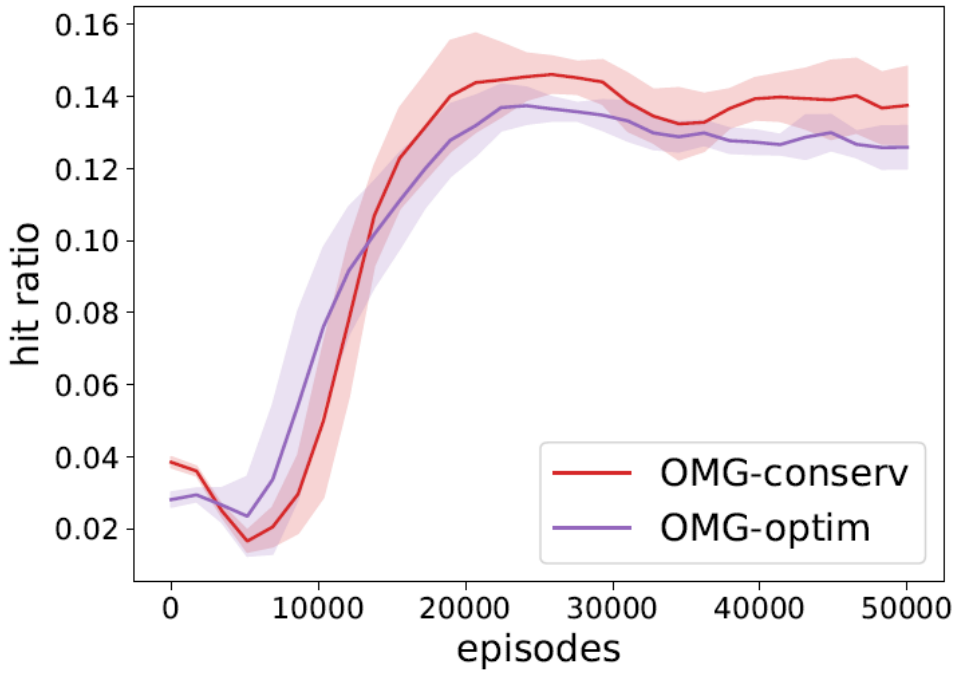
Ablation Study



Subgoal selection strategies



Experiment



Hit ratio of subgoal state

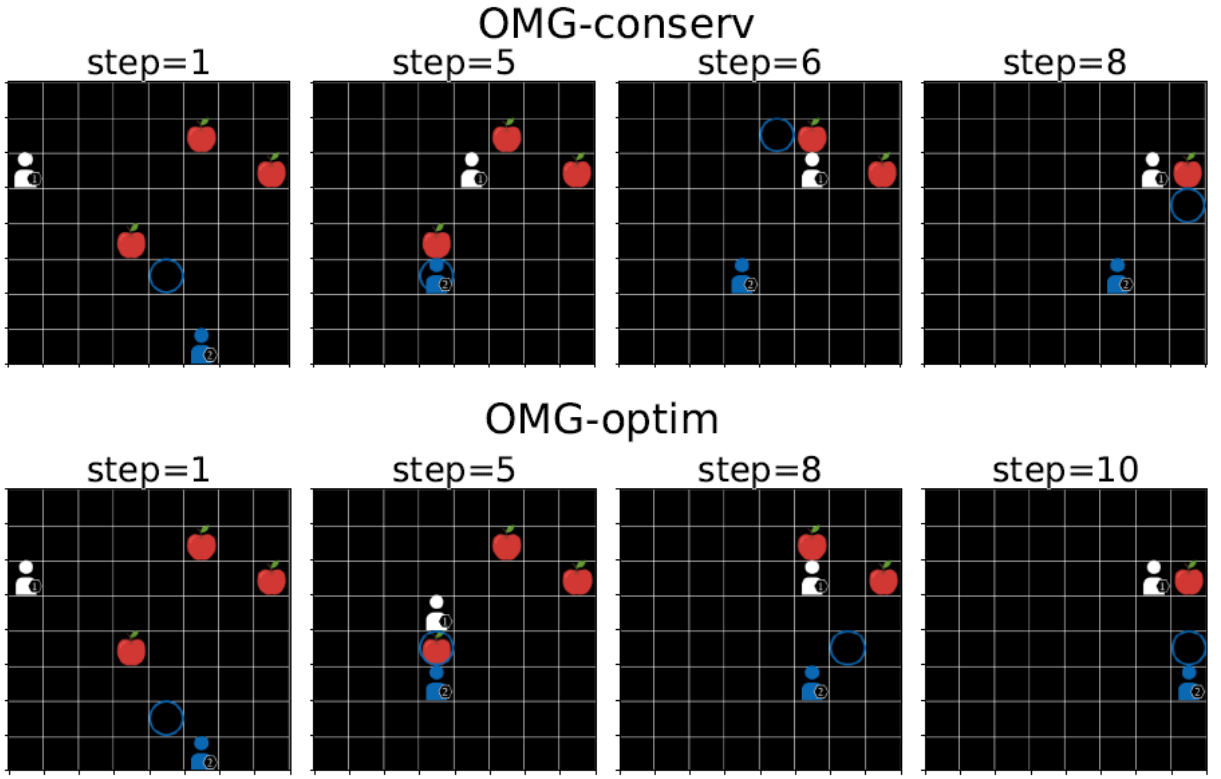


Illustration of inferred subgoal state

Summary

01 Offers a novel perspective on opponent modeling by focusing on subgoal inference.

02 Intuition is consistent with the way of thinking in human being.

03 The experiment results show that OMG get state-of-the-art in this domain.

Thank you for listening.