

# Fuzzy Rule BayesOpt AutoRL

Bayesian Optimization

$$z^* = \operatorname{argmax} f(z)$$

f: Team 1's End-Game Winrate/Reward/Advantage

$z = [$ 
Fuzz Fuzz Fuzz Fuzz Fuzz Fuzz  $]$   
 $z = [ 1.00, 1.00, 4.00, 6.00, 0.42, ... ]$

Fuzzy Logic Decoder

Fuzzy Logic Encoder



Fuzzy Rules



Opponent  
RL Learner 3

Opponent  
RL Learner 2

POV (Point of View) of Learner 1



- [P] Recent Winrate
- [P] Accumulated Winrate
- [P] ...
- [P] ...
- [P] Agent Life Count
- [P] Number of Alive Agent
- [P] ...
- [P] ...

agent trajectories  
(observation & reward)

agent policies (actions)



Mamdani-type Fuzzy Controller (MFC)



Fuzzy Feedback Adjustment

Model Feedback Loop



Improved RL Learner for Team 1

RL Learner  
Dynamic  
Adjustment

- AC Hyper-Parameters
- Actor's Rewards
- Exploration-Exploitation balance
- ...