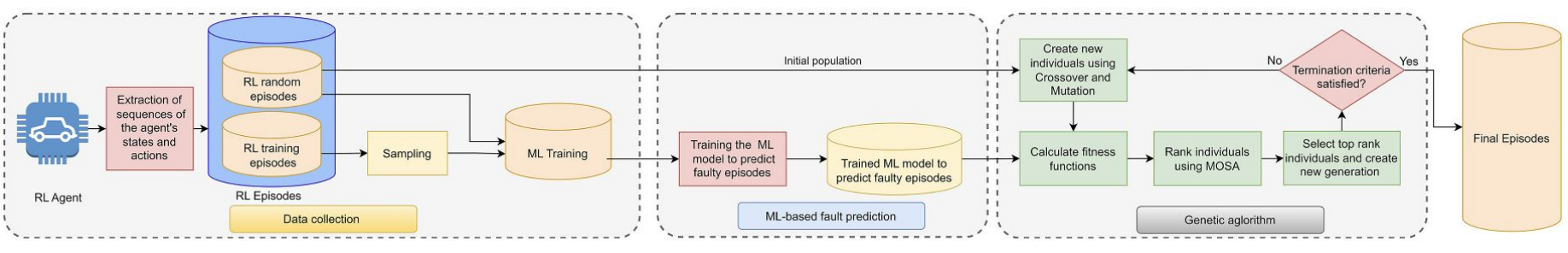
Systematic Literature Review of Testing Tools and Techniques for Reinforcement Learning Agents

Groza Iulia, Havirneanu Andrei, Ilie Andreea



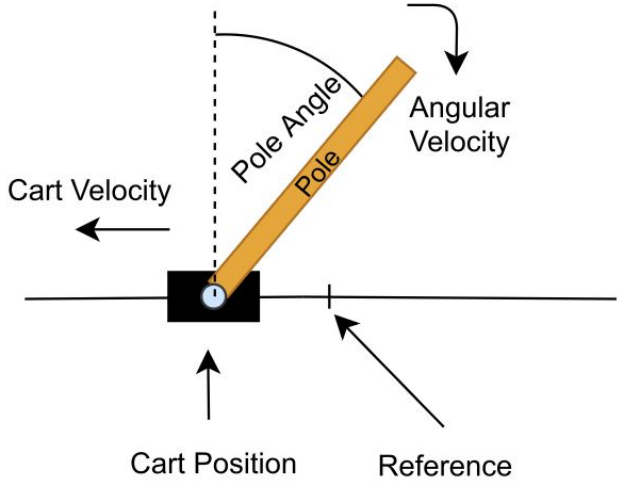


Fig. 3. *Cart-Pole* balancing problem.

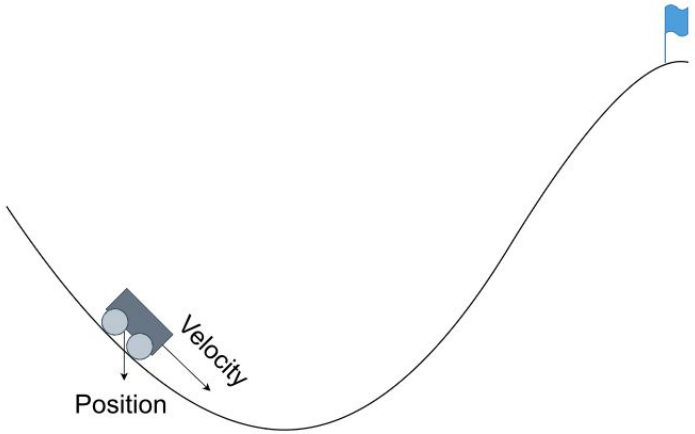
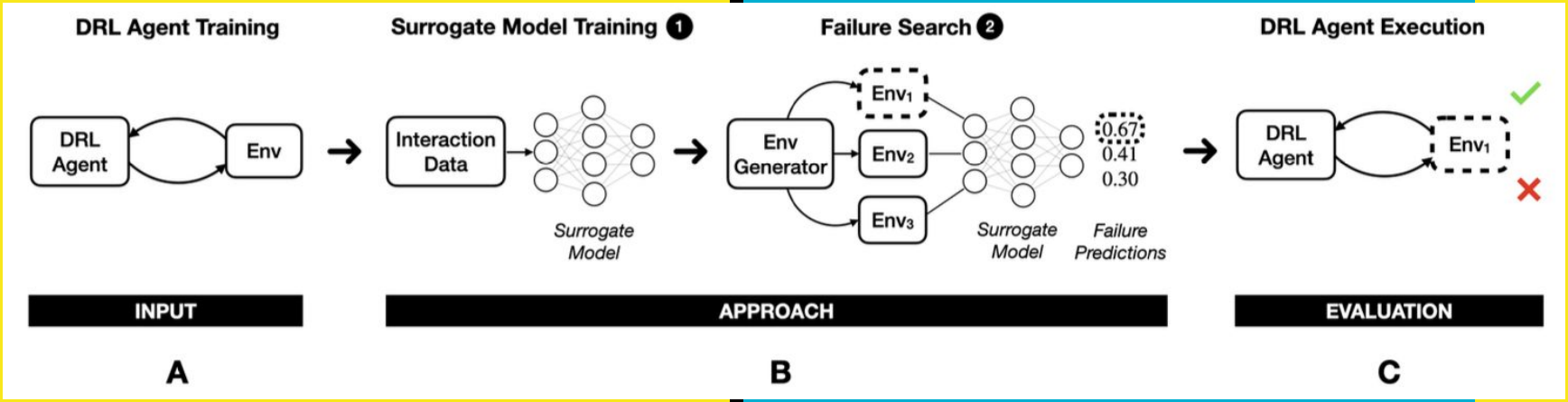
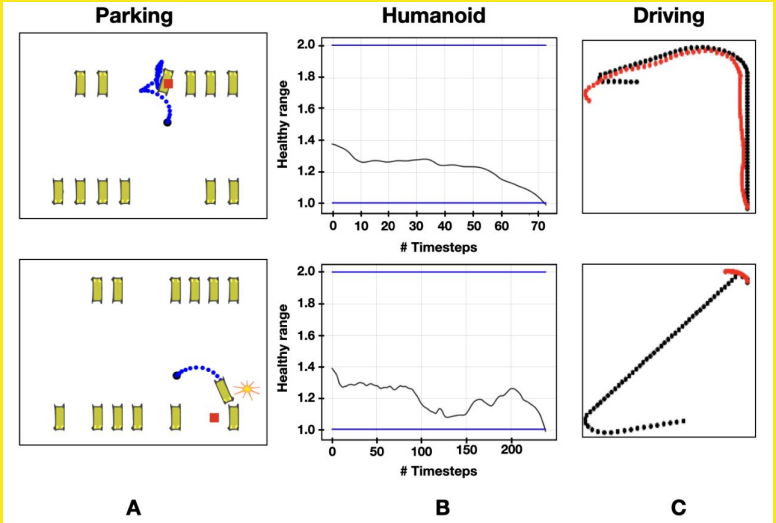
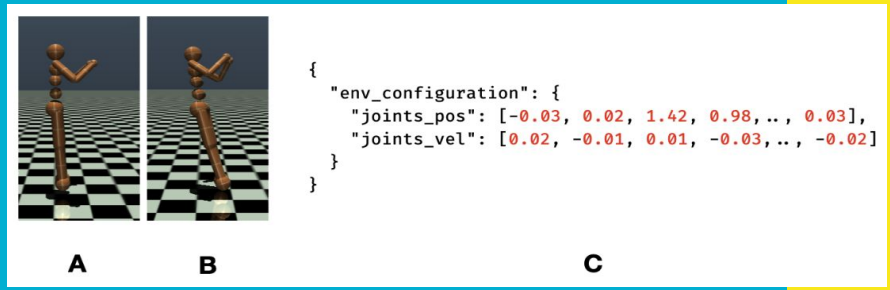
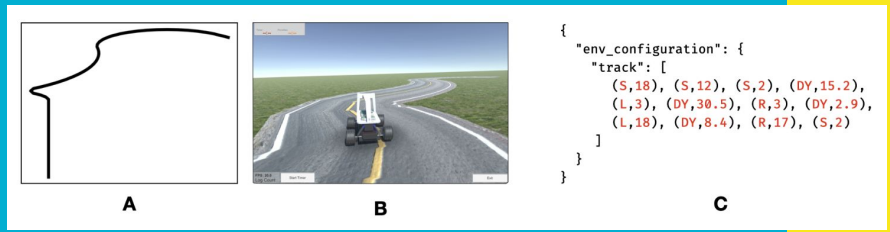


Fig. 4. *Mountain Car* problem.


DRL Agent Training
Surrogate Model Training 1
Failure Search 2
DRL Agent Execution
DRL Agent
Env
Interaction Data
Surrogate Model
Env Generator
Env1
Env2
Env3
0.67
0.41
0.30
Surrogate Model
Failure Predictions
DRL Agent
Env1
INPUT
APPROACH
EVALUATION
A
B
C

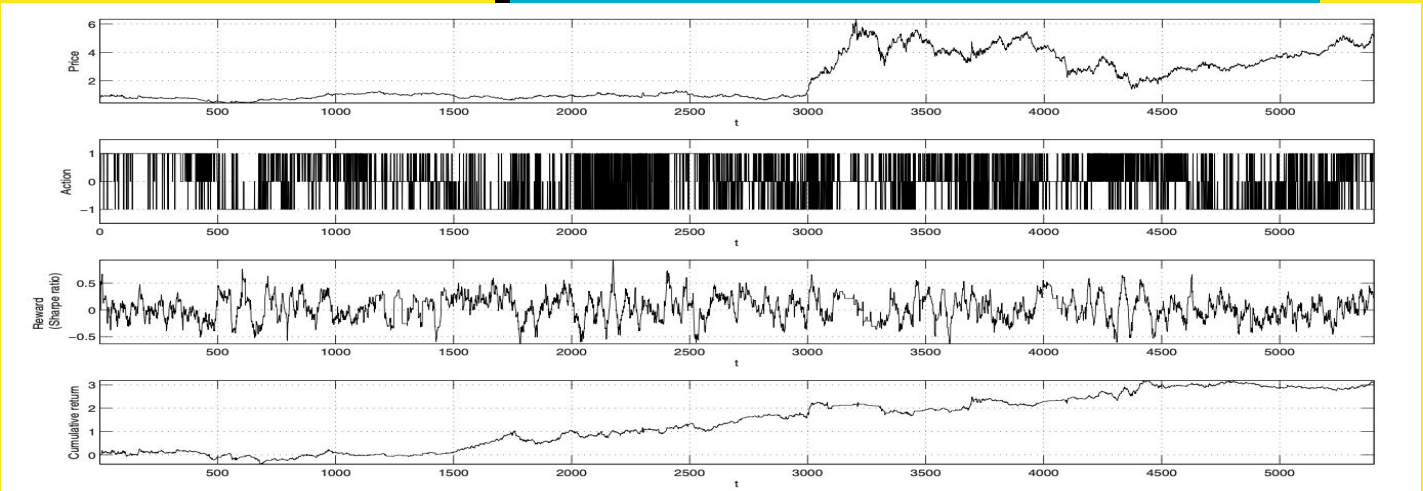

Parking
Humanoid
Driving
Healthy range
Timesteps
A
B
C

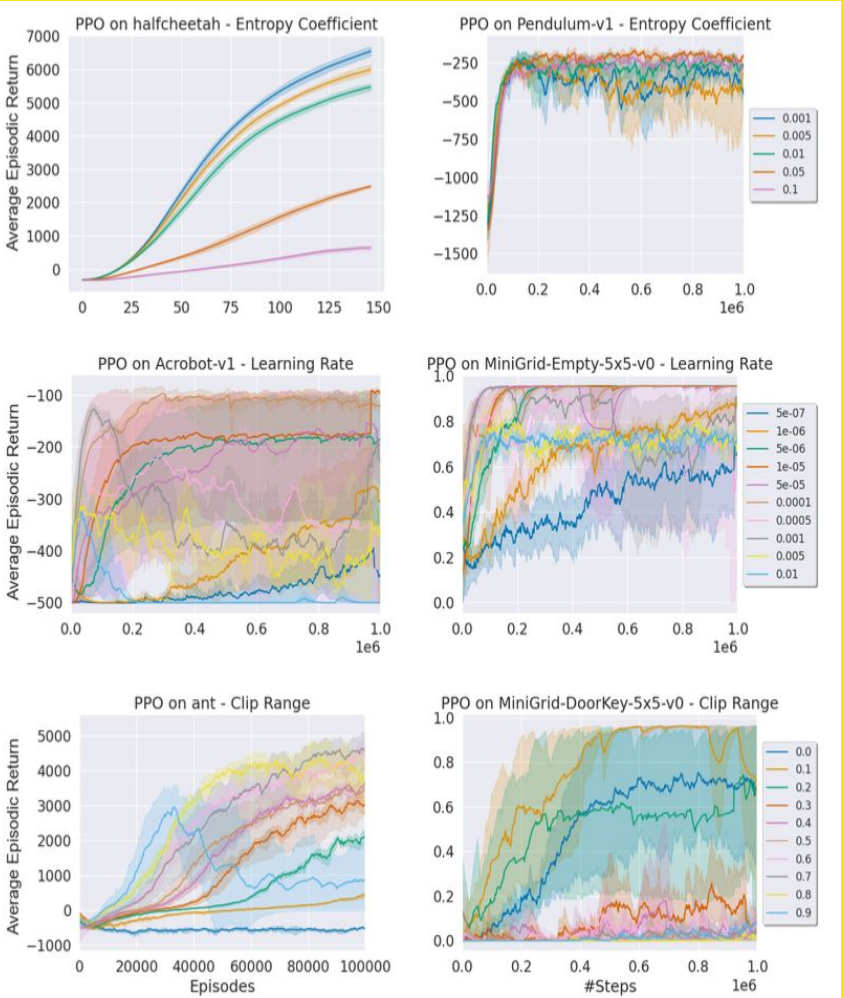

{
"env_configuration": {
"joints_pos": [-0.03, 0.02, 1.42, 0.98, .., 0.03],
"joints_vel": [0.02, -0.01, 0.01, -0.03, .., -0.02]
}
}
A
B
C


{
"env_configuration": {
"track": [
(S,18), (S,12), (S,2), (DY,15.2),
(L,3), (DY,30.5), (R,3), (DY,2.9),
(L,18), (DY,8.4), (R,17), (S,2)
]
}
}
A
B
C

Approach	L	Statistics	Artificial time series	Real time series
QLa	5	μ	472.02%	40.00%
		σ	32.79%	139.09%
		Confidence interval	[407.76%, 536.29%]	[−226.73%, 306.73%]
KebRL	5	μ	435.42%	−7.99%
		σ	41.13%	131.26%
		Confidence interval	[354.81%, 516.02%]	[−265.26%, 249.28%]
QLa	22	μ	337.68%	92.92%
		σ	40.80%	149.28%
		Confidence interval	[257.71%, 417.64%]	[−199.66%, 385.50%]
KbRL	22	μ	216.61%	13.20%
		σ	49.98%	153.34%
		Confidence interval	[118.64%, 314.58%]	[−287.34%, 313.74%]

Table 1. Statistics about the final cumulative returns.





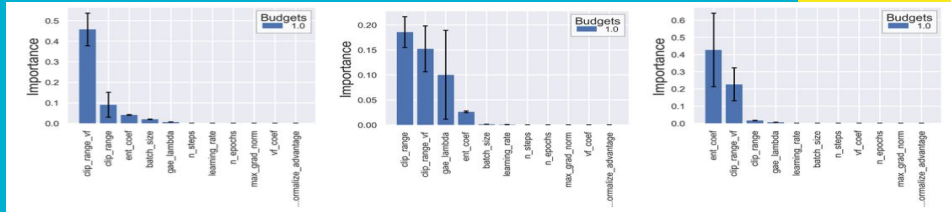


Figure 27: PPO Hyperparameter Importances on Brax Ant (left), Halfcheetah (middle) and Humanoid (right).

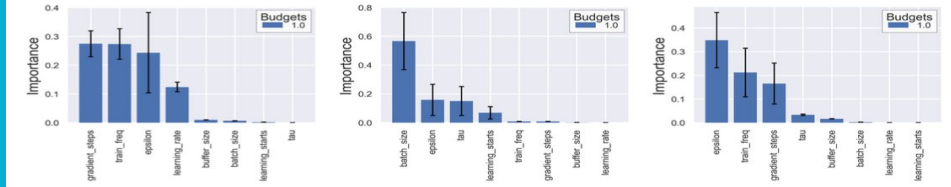


Figure 28: DQN Hyperparameter Importances on Acrobot (left), MiniGrid Empty (middle) and MiniGrid DoorKey (right).

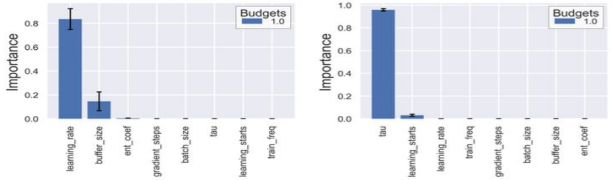


Figure 29: SAC Hyperparameter Importances on Pendulum (left) and Brax Ant(right).

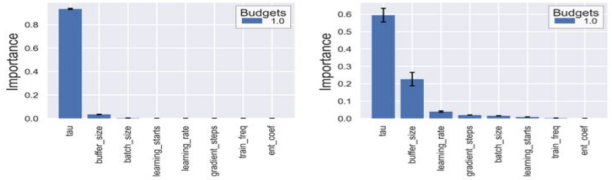


Figure 30: SAC Hyperparameter Importances on Brax Halfcheetah (left) and Humanoid (right).

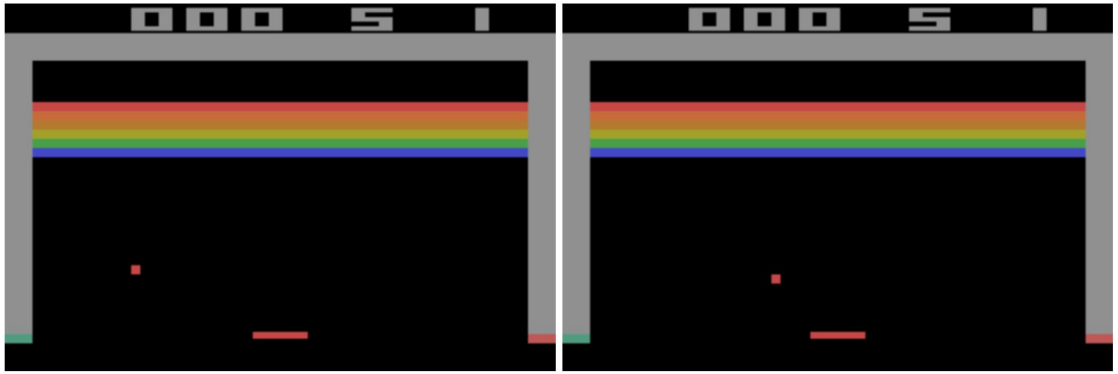
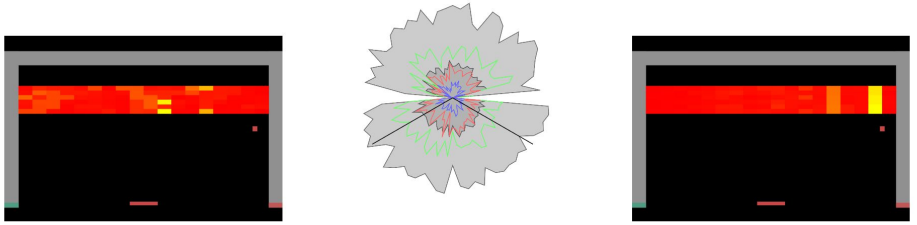
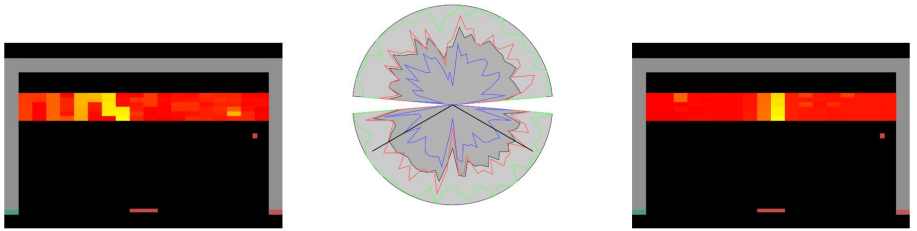


Figure 1: (*Left*: ALE. *Right*: TOYBOX.) Images of near-start frames for both Atari and TOYBOX implementations of Breakout.



(a) Brick Elimination (1e7) (b) Start Angle Invariance (1e7) (c) Tunnel Exploitation (1e7)



(d) Brick Elimination (5e7) (e) Start Angle Invariance (5e7) (f) Tunnel Exploitation (5e7)

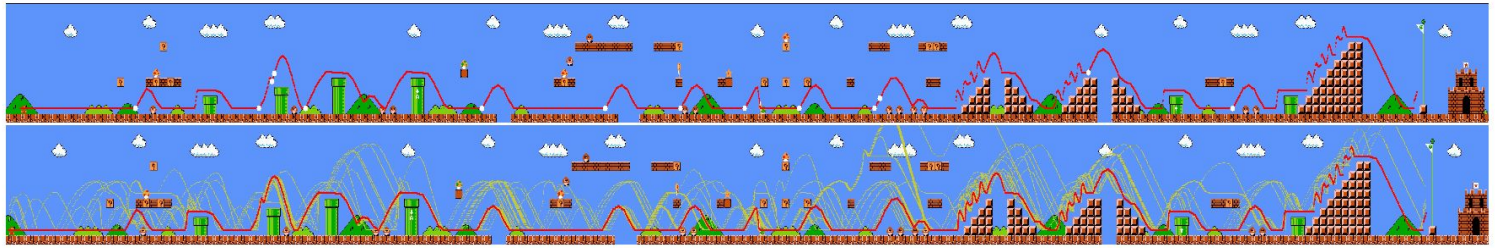


Figure 1: Super Mario Bros. Up: Reference Trace and Boundary States. Down: Reference Trace and Fuzz Traces.

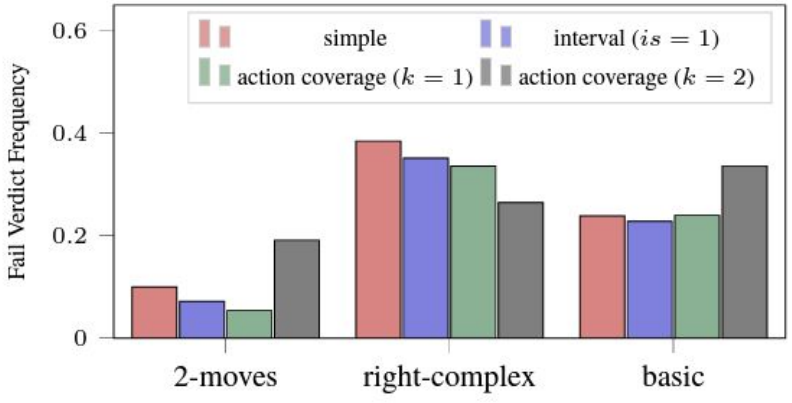
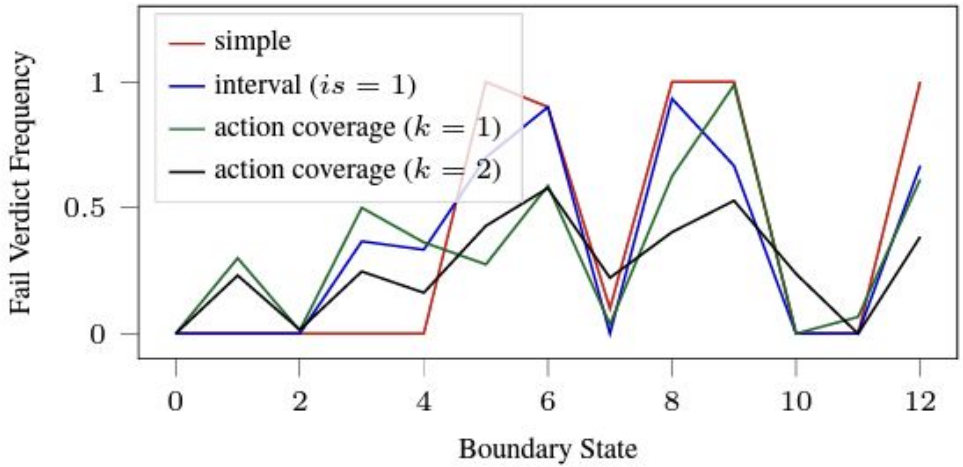
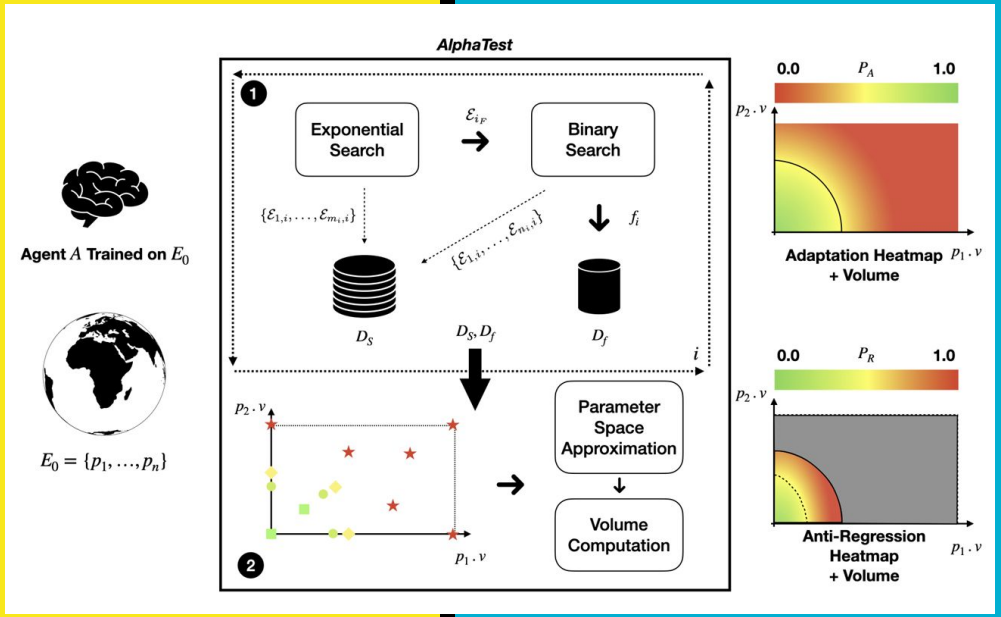
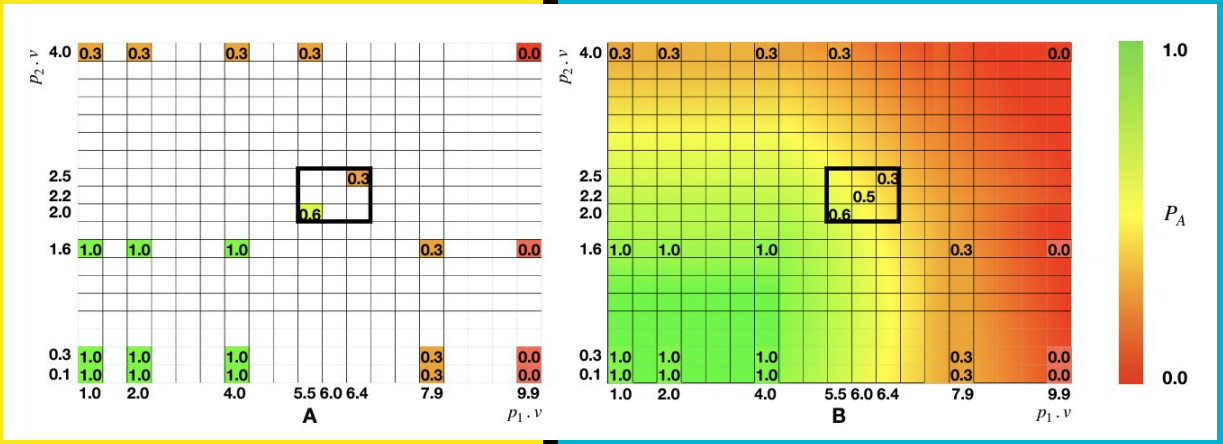


Figure 3: Safety Testing: Relative frequency of fail verdicts




AlphaTest
Agent A Trained on E_0
$E_0 = \{p_1, \ldots, p_n\}$
1
Exponential Search
$\mathcal{E}_{iF}$
Binary Search
$\{\mathcal{E}_{1,i}, \ldots, \mathcal{E}_{m_i,i}\}$
$\{\mathcal{E}_{1,i}, \ldots, \mathcal{E}_{m_i,i}\}$
f_i
D_S
D_S, D_f
D_f
i
$p_2 \cdot v$
$p_1 \cdot v$
2
Parameter Space Approximation
Volume Computation
0.0
P_A
1.0
$p_2 \cdot v$
Adaptation Heatmap + Volume
$p_1 \cdot v$
0.0
P_R
1.0
$p_2 \cdot v$
Anti-Regression Heatmap + Volume
$p_1 \cdot v$


$p_2 \cdot v$
4.0
2.5
2.2
2.0
1.6
0.3
0.1
1.0
2.0
4.0
5.5 6.0 6.4
7.9
9.9
$p_1 \cdot v$
A
B
1.0
P_A
0.0

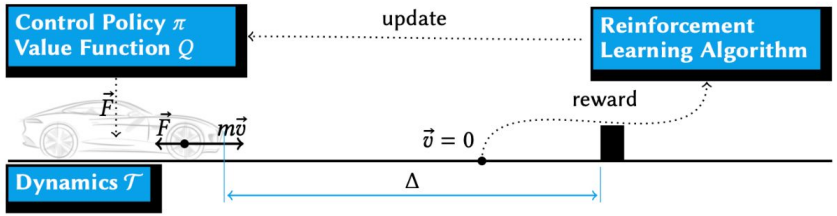


Fig. 1. Example: A car moves with velocity $\vec{v}$ towards a fixed obstacle at distance Δ, learning how to brake. The control policy chooses a deceleration with which to brake. The agent receives a reward based on the location where stopped and updates the policy

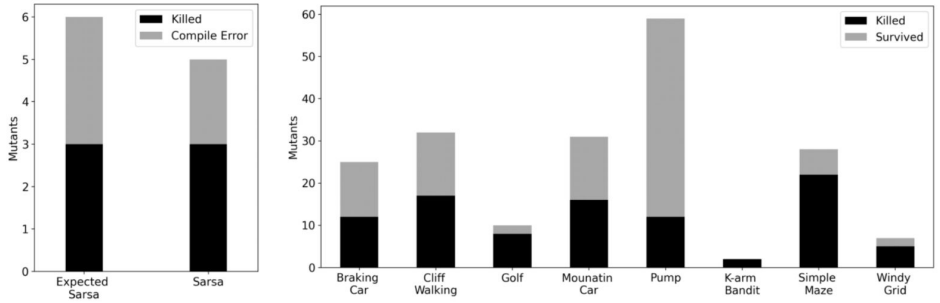


Fig. 8. Mutation results with generic tests for SARSA/Expected SARSA (left) and case studies (right).

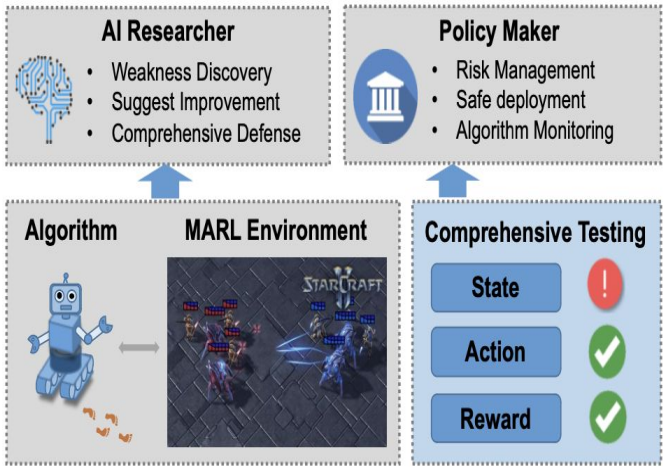
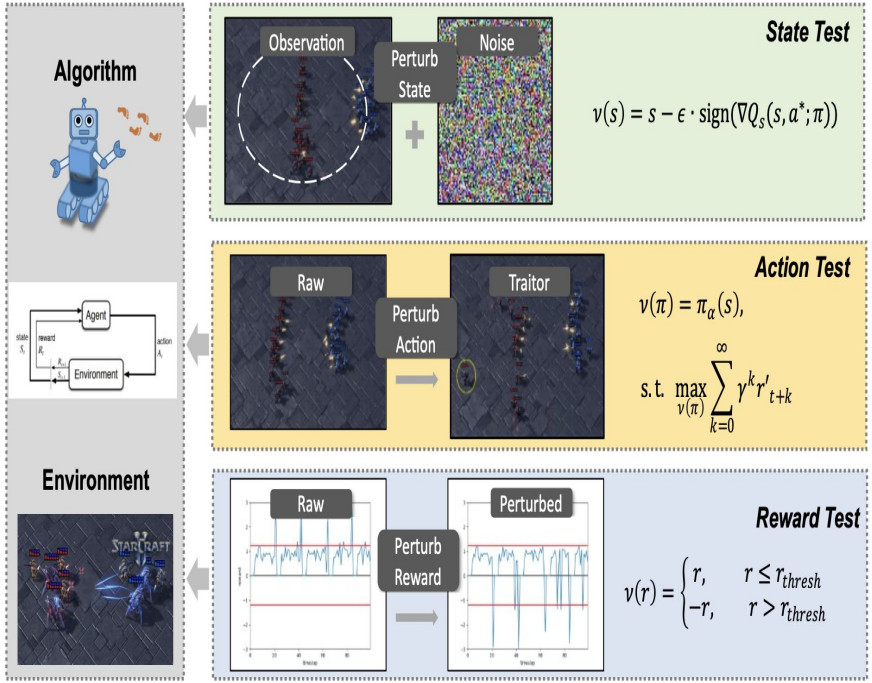


Figure 1. An overview of our work. c-MARL algorithm could be attacked from three aspects, namely state, action and reward. We test the robustness of c-MARL from these aspects.



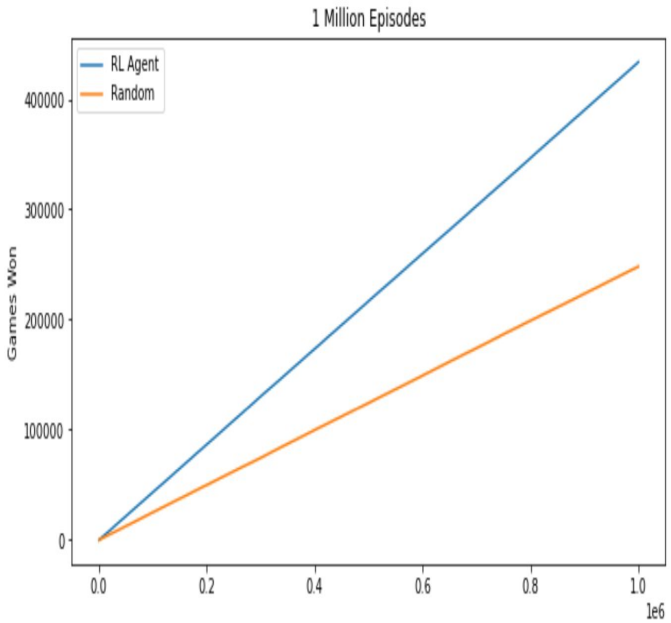


Figure 1: Performance of RL Agent vs Random Agent

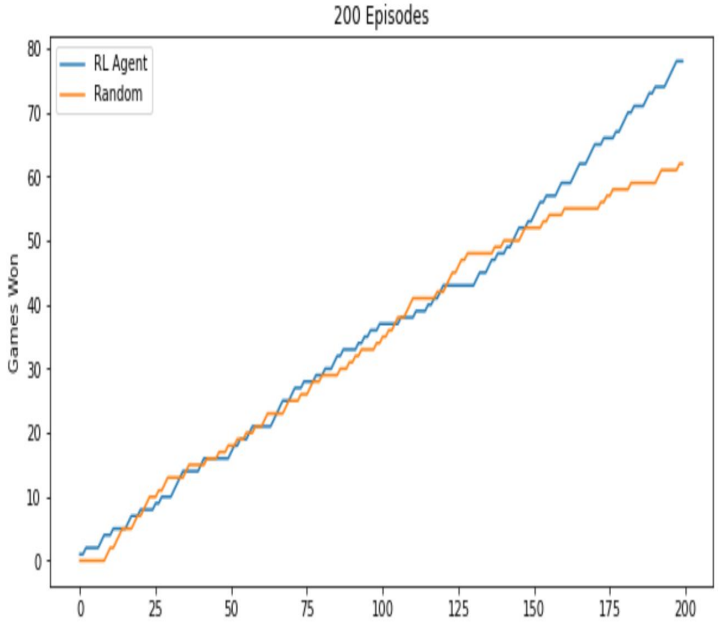


Figure 2: First 200 games of RL Agent vs Random Agent

References:

1. “A Search-Based Testing Approach for Deep Reinforcement Learning Agents”, Zolfagharian, Amirhossein and Abdellatif, Manel and Briand, 2023
2. “Testing of Deep Reinforcement Learning Agents with Surrogate Models”, Biagiola, Matteo and Tonella, Paolo, 2024
3. “Testing Different Reinforcement Learning Configurations for Financial Trading: Introduction and Applications”, Francesco Bertoluzzo and Marco Corazza, 2012
4. “Hyperparameters in Reinforcement Learning and How To Tune Them”, Eimer, Theresa and Lindauer, Marius and Raileanu, Roberta, 2023
5. “TOYBOX: Better Atari Environments for Testing Reinforcement Learning Agents”, Foley, John and Tosch, Emma and Clary, Kaleigh and Jensen, David, 2019
6. “Search-Based Testing of Reinforcement Learning”, Martin Tappler and Filip Cano Córdoba and Bernhard K. Aichernig and Bettina Könighofer, 2022
7. “Testing the Plasticity of Reinforcement Learning Based Systems” by Biagiola, Matteo and Tonella, Paolo, 2022
8. “Formal Specification and Testing for Reinforcement Learning", Mahsa Varshosaz and Mohsen Ghaffari and Einar Broch Johnsen and Andrzej Wąsowski, 2023
9. “Towards Comprehensive Testing on the Robustness of Cooperative Multi-agent Reinforcement Learning”, Jun Guo and Yonghong Chen and Yihang Hao and Zixin Yin and Yin Yu and Simin Li, 2022
10. “Reinforcement Learning Agents in Colonel Blotto”, Joseph Christian G. Noel, 2022