Attention-based Curiosity-driven Exploration in Deep Reinforcement Learning

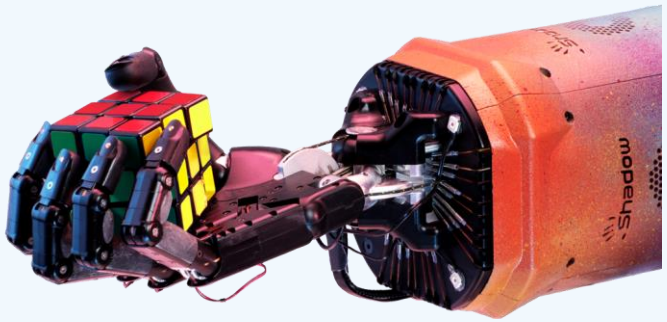
Authors: Patrik Reizinger, Márton Szemenyei

Supported by the Higher Education Excellence Program in the frame of Artificial Intelligence research area of BUTE (BME FIKP-MI/FM), ÚNKP-19-2 New National Excellence Program of the Ministry for Innovation and Technology and by the EFOP 3.6.1.-16-2016-00014 program

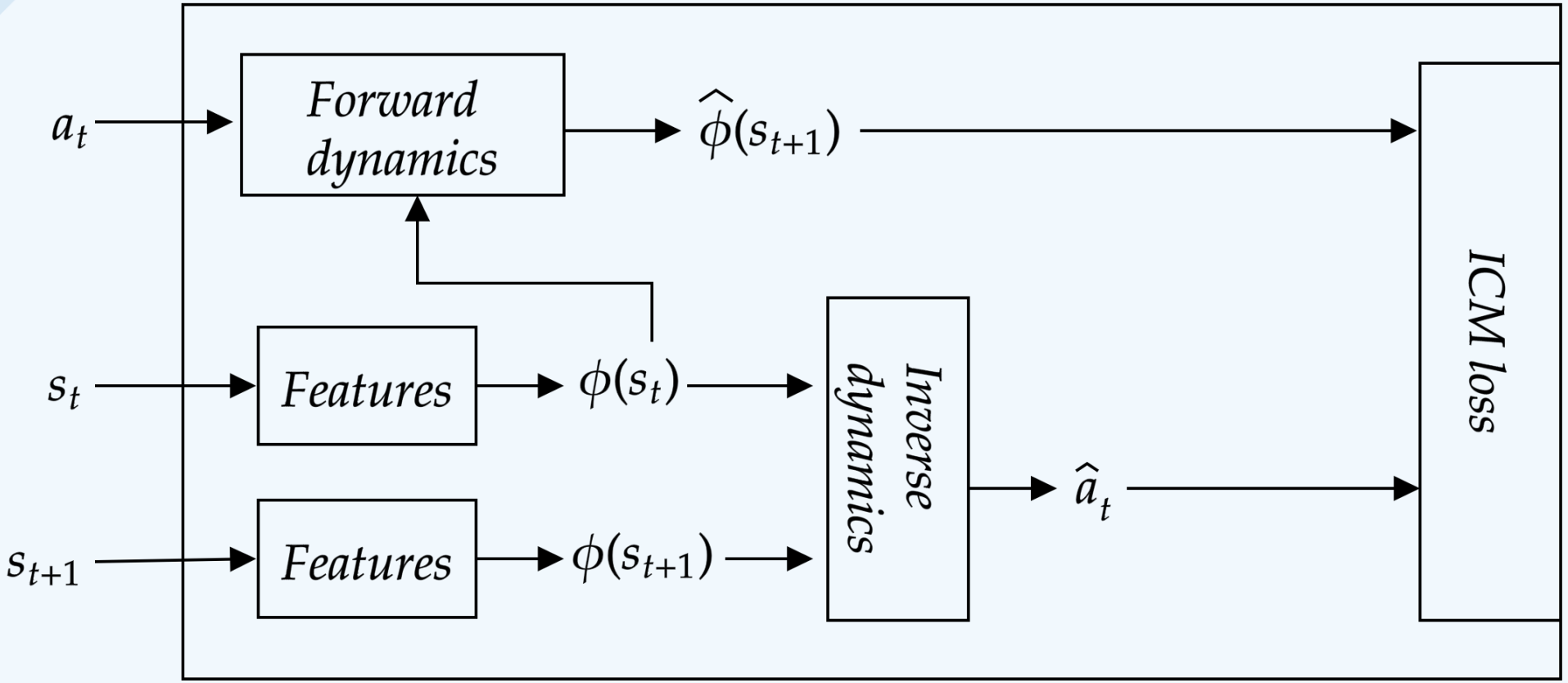
08.-04.05.2020

Motivation

- Develop feature-selective intrinsic reward formulations
- Apply curiosity/attention in an adaptive manner
- Extend the applicability of previous work

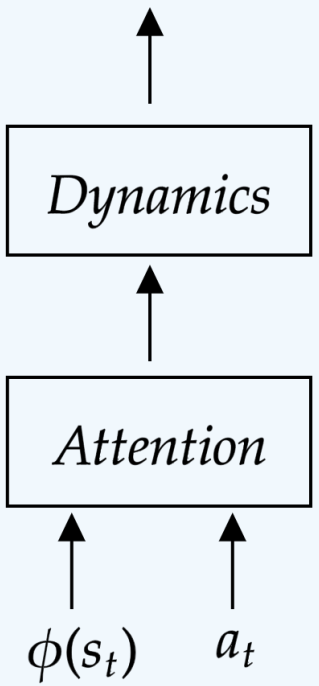


Intrinsic Curiosity Module (ICM)

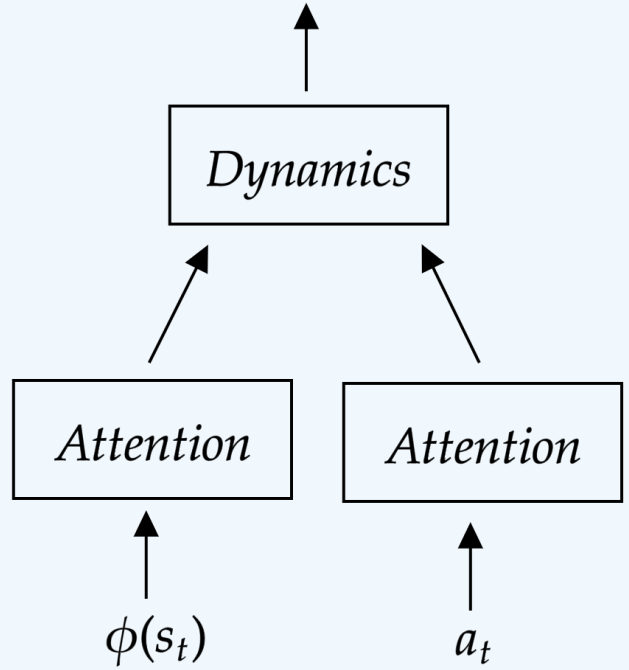


ICM & Attention mechanism

Single attention

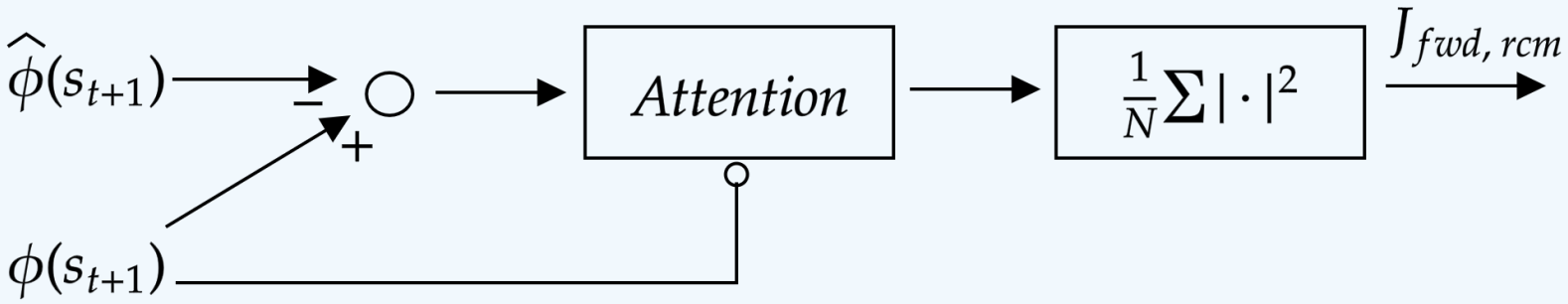


Double attention

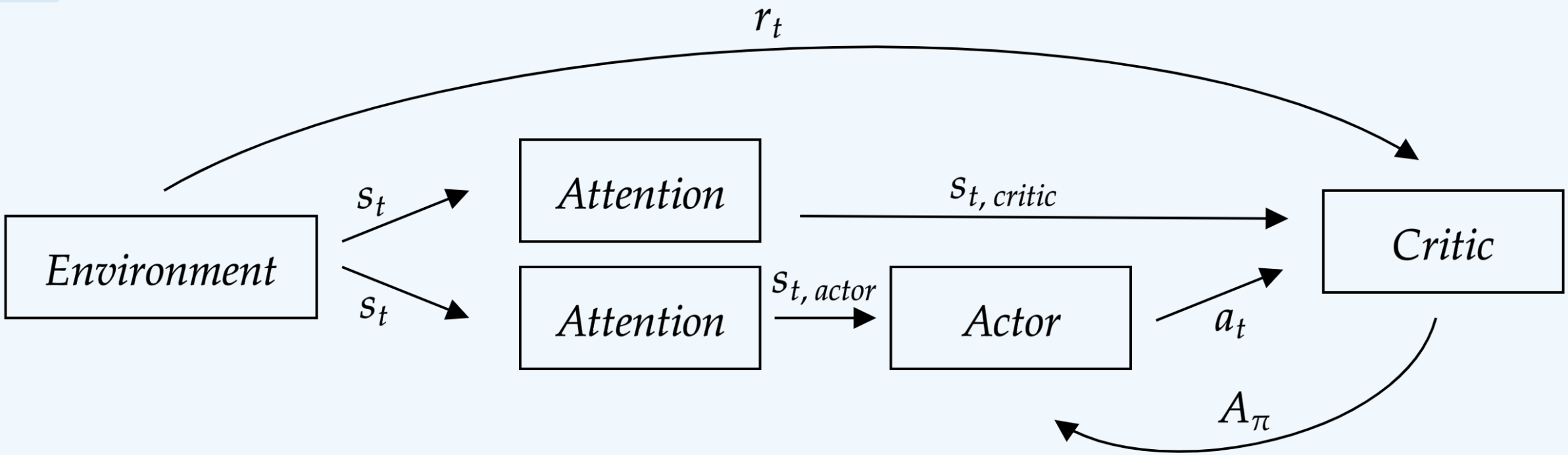


Rational Curiosity Module (RCM)

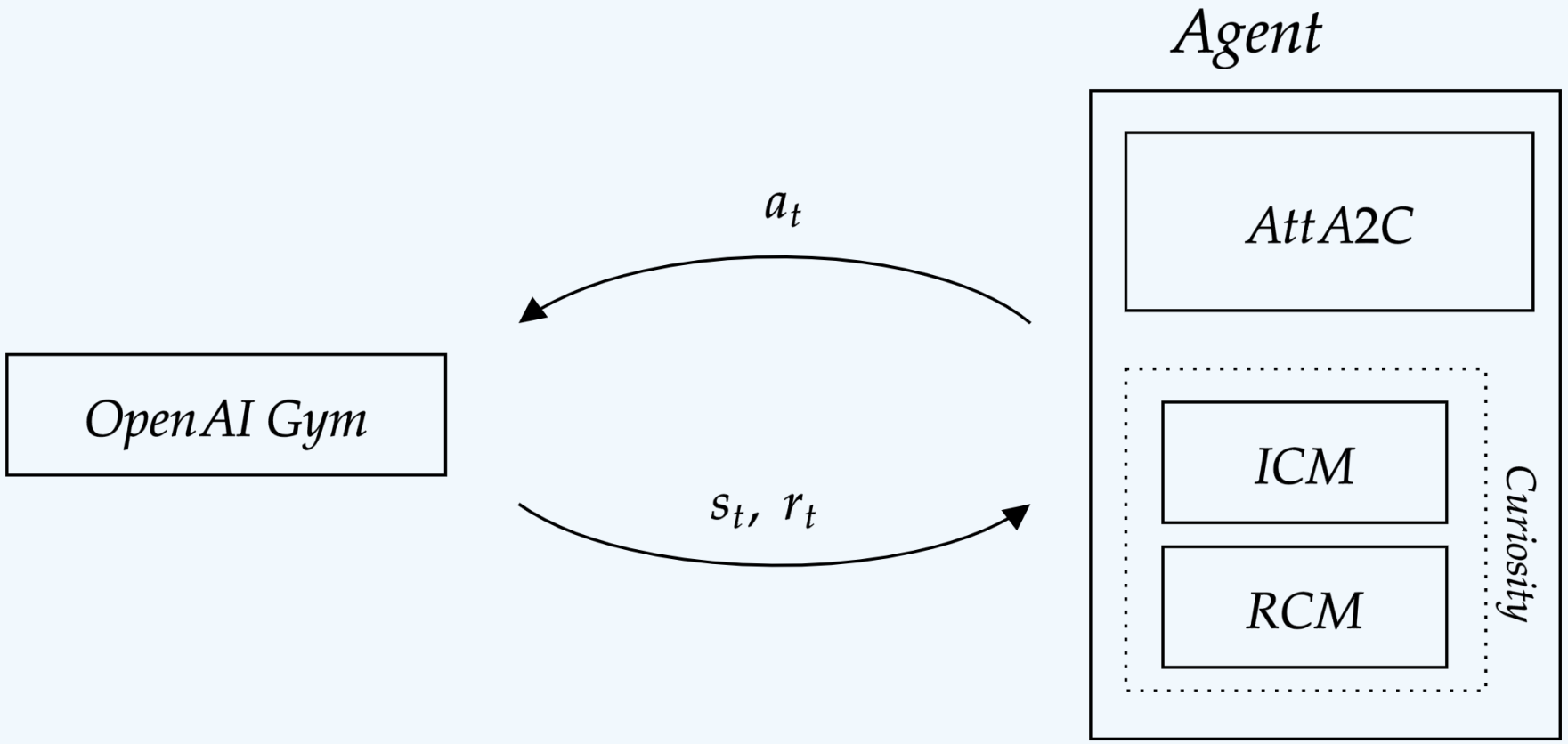
- Action- and state-selective weighting of the curiosity loss
- Analog to Adaptive Weighted Least Squares



AttA2C

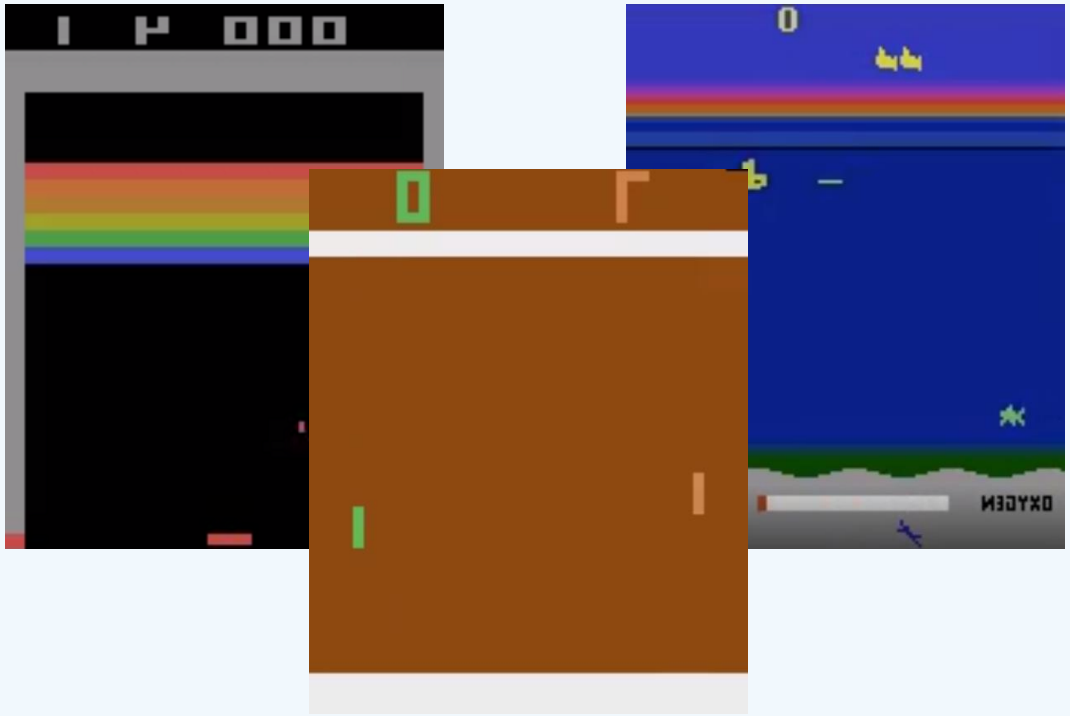

r_t
Environment
s_t
s_t
Attention
Attention
$s_{t,\,critic}$
$s_{t,\,actor}$
Actor
Critic
a_t
A_π

Implementation



Experiments

Environments

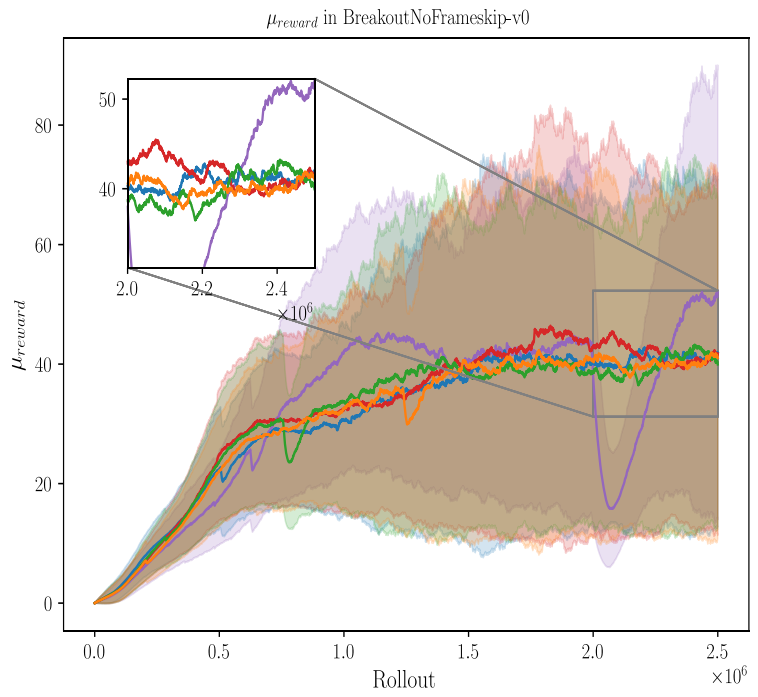


Parameters

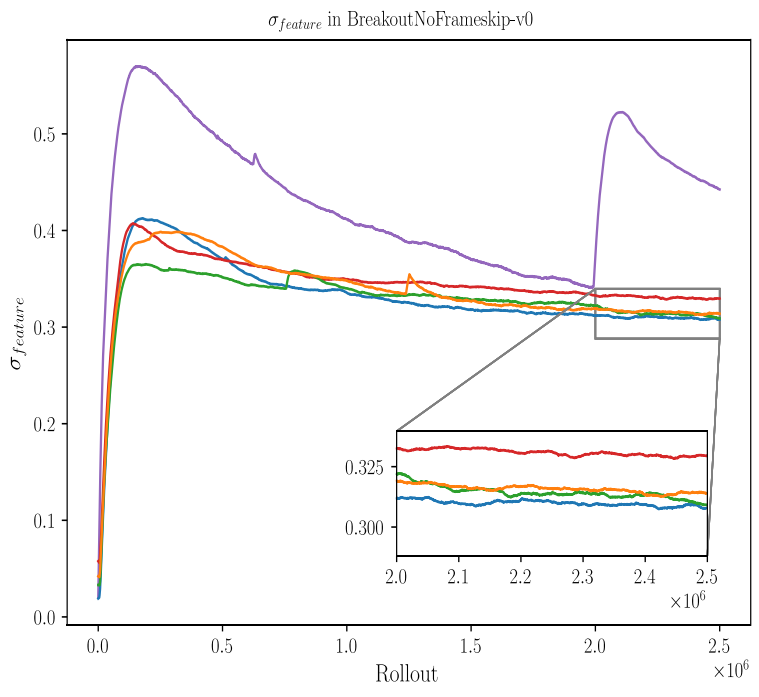
- Deterministic & stochastic variants
- 4 parallel environments
- 2.5 million steps (5-5 frames)
- 5 different agents

Results – Breakout

Mean reward



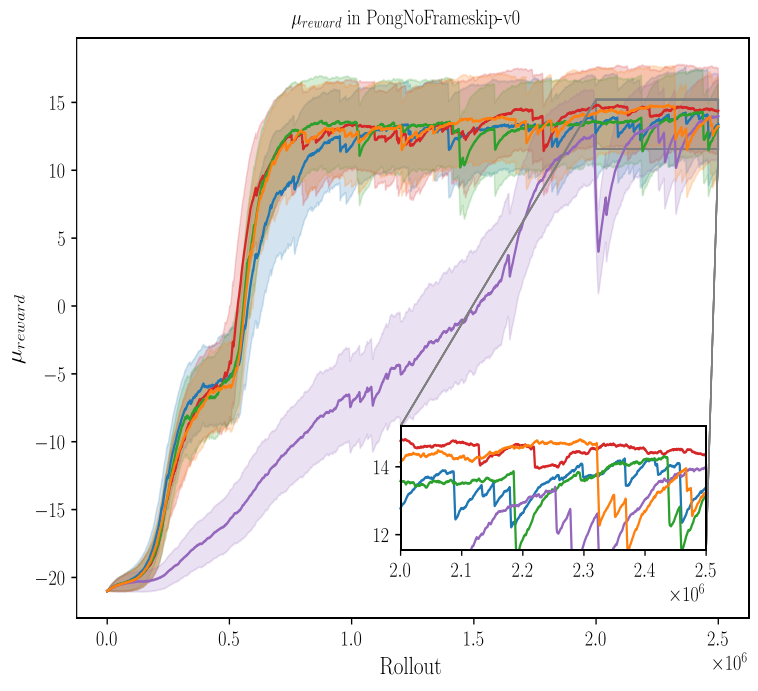
Features



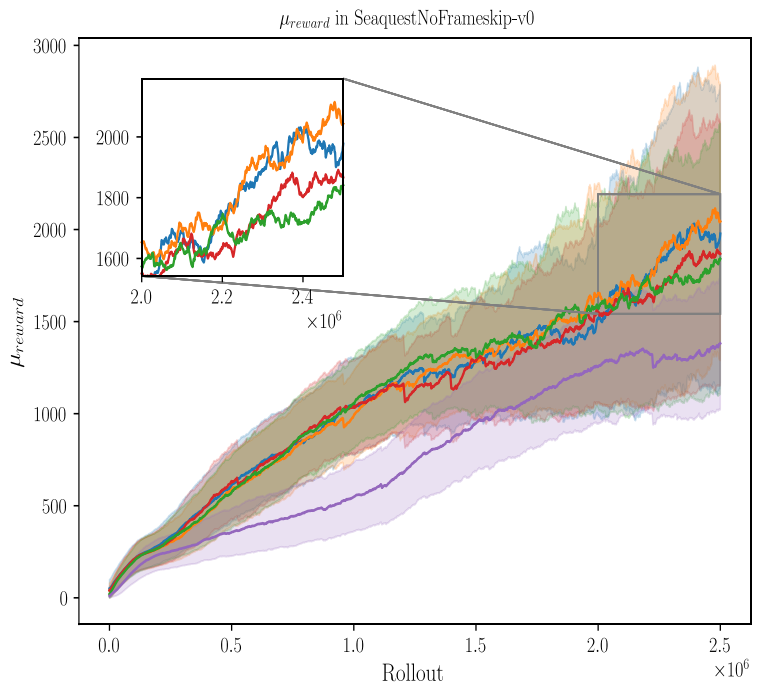
ICM AttA2C Single Attention Double Attention RCM

Results – Pong & Seaquest

Pong



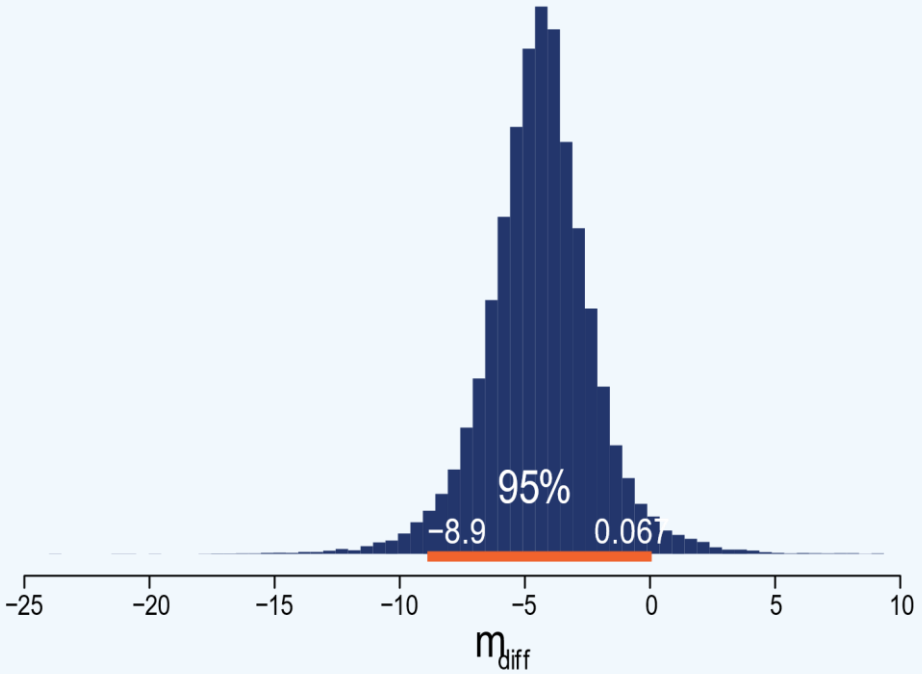
Seaquest



ICM AttA2C Single Attention Double Attention RCM

Results comparison

Significance (RCM)



Methodology

- Normalized reward
- Bayesian paired t-test
- RCM: 97%, Single Attention: 92% confidence

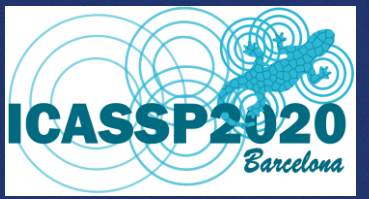
Conclusion & Future Work

- Successful application of attention mechanism
- RCM: statistically siginificant improvement
- Long term generalization
- Weighting intrinsic & extrinsic rewards

https://github.com/rpatrik96/AttA2C

Attention-based Curiosity-driven Exploration in Deep Reinforcement Learning

Authors: Patrik Reizinger (rpatrik96@sch.bme.hu), Márton Szemenyei



Supported by the Higher Education Excellence Program in the frame of Artificial Intelligence research area of BUTE (BME FIKP-MI/FM), ÚNKP-19-2 New National Excellence Program of the Ministry for Innovation and Technology and by the EFOP 3.6.1.-16-2016-00014 program

08.-04.05.2020