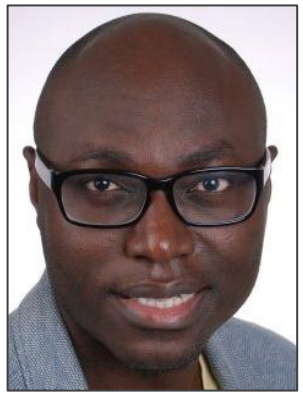
Deep Reinforcement Learning on a Budget
3D Control and reasoning without a supercomputer

ICPR 2020

Webpage: https://edbeeching.github.io/papers/3d_control_deep_rl

Edward Beeching

Jilles Dibangoye

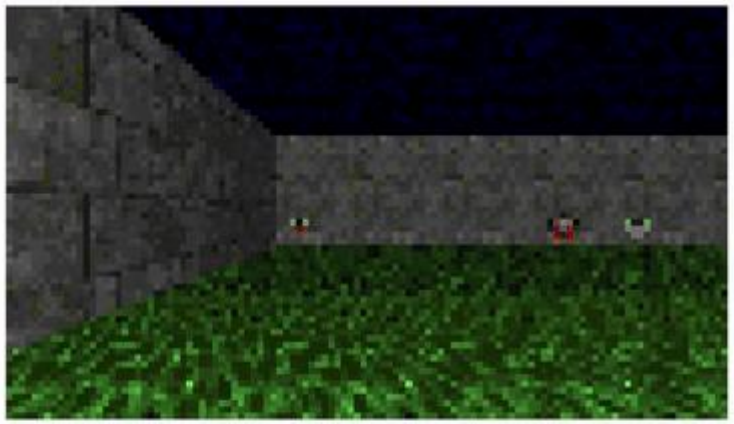
Olivier Simonin

Christian Wolf

Introduction

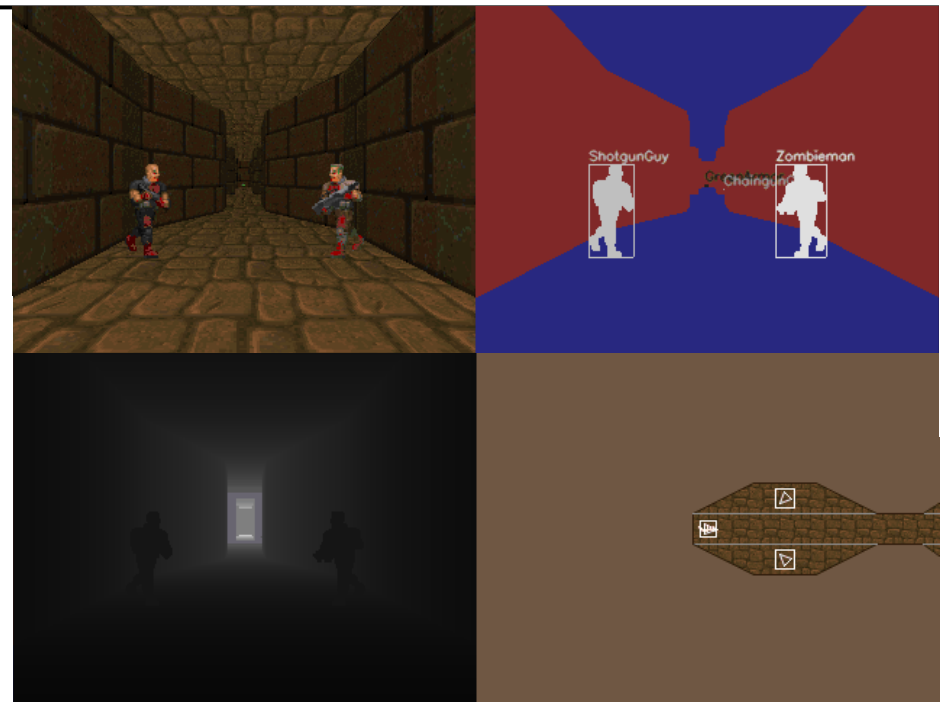
- We provide challenging scenarios for Deep Reinforcement Learning agents in partially observable 3D environments.

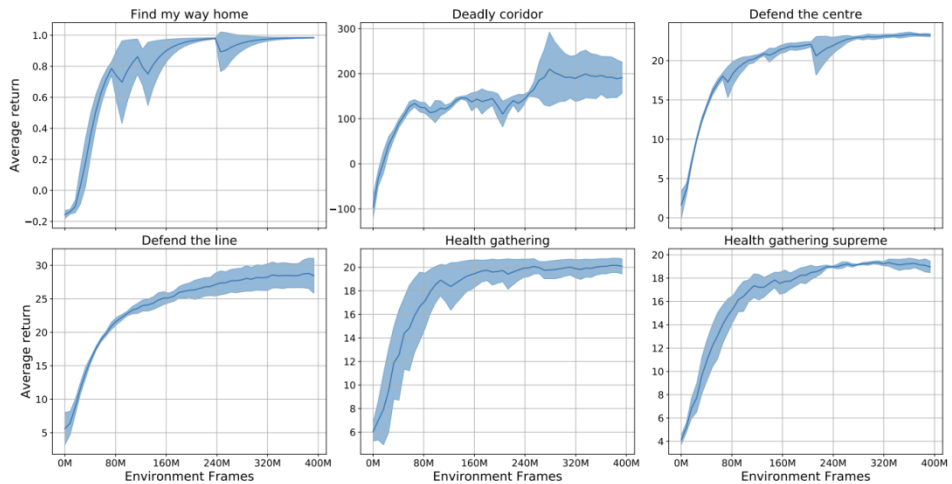
- This work:
  - Extends the ViZDoom simulator with 4 new scenarios
  - Trains and evaluates a recurrent Deep RL agent to identify its limitations and areas for improvement
  - Evaluates the generalization performance in unseen environment configurations.
- Our scenarios:
  - Act as a proxy to real world environments
  - Test reasoning, memory, generalization and planning in complex 3D environments.
  - Provide a standard benchmark for Deep RL algorithms that can be trained in under 24 hours



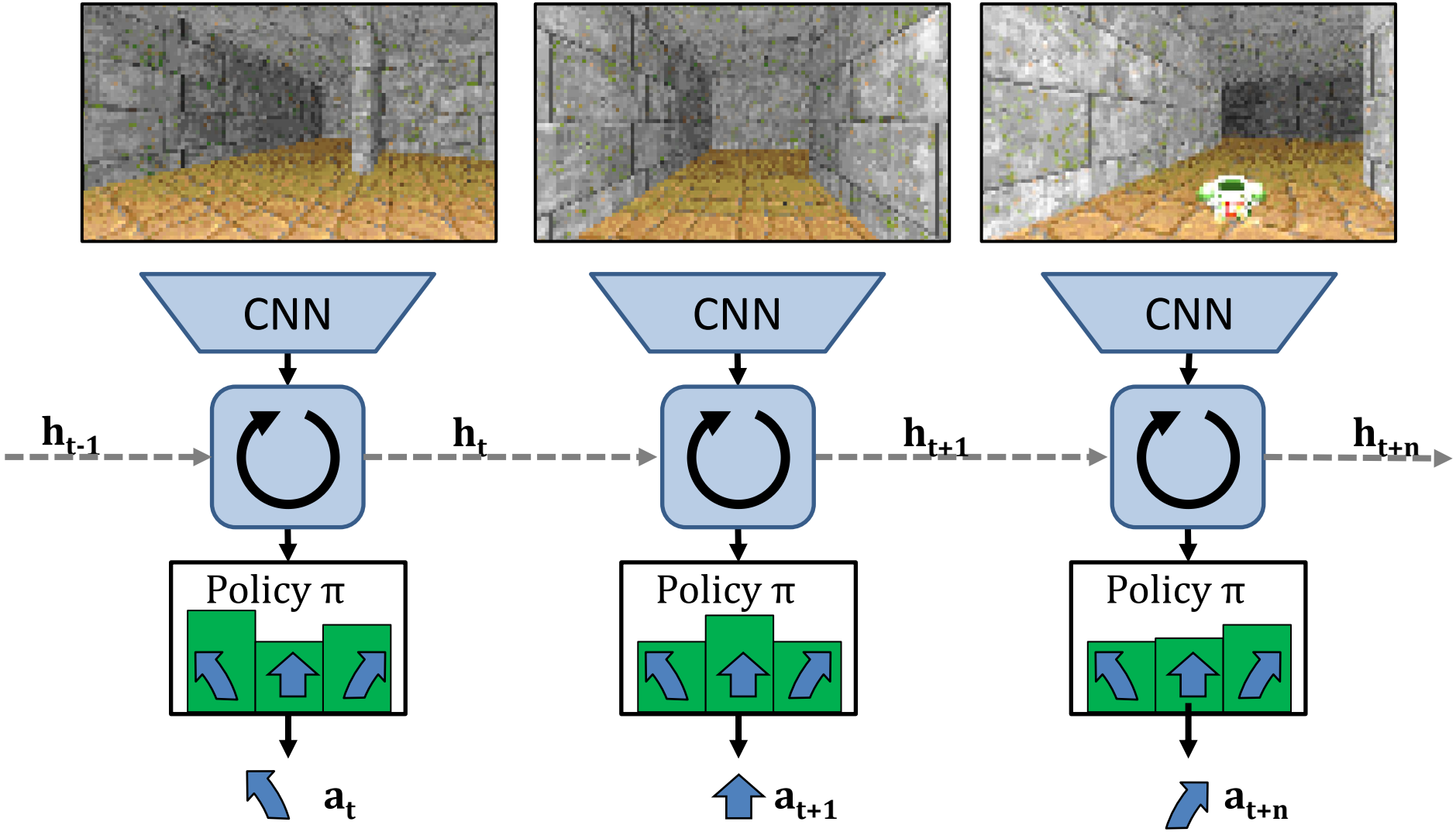
Motivation

- ViZDoom (Wydmuch et al., 2016)
  - Highly optimized software
  - Runs at 100x real time on modern computers
  - Mature set of tools for the creation of new scenarios
  - Includes some simple scenarios
- Lacks:
  - Challenging scenarios
  - Scenarios that test generalization in unseen environments
- Our aim:
  - To create a set of memory based tasks that can act as proxies for real world robotic navigation and planning problems.
  - Evaluate the performance of Deep RL agents





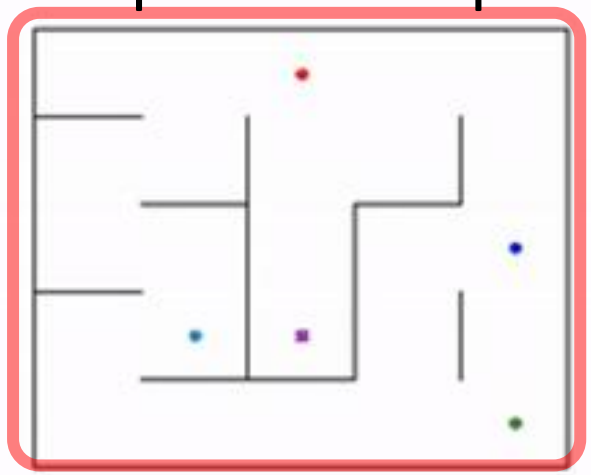
Benchmark recurrent agent architecture



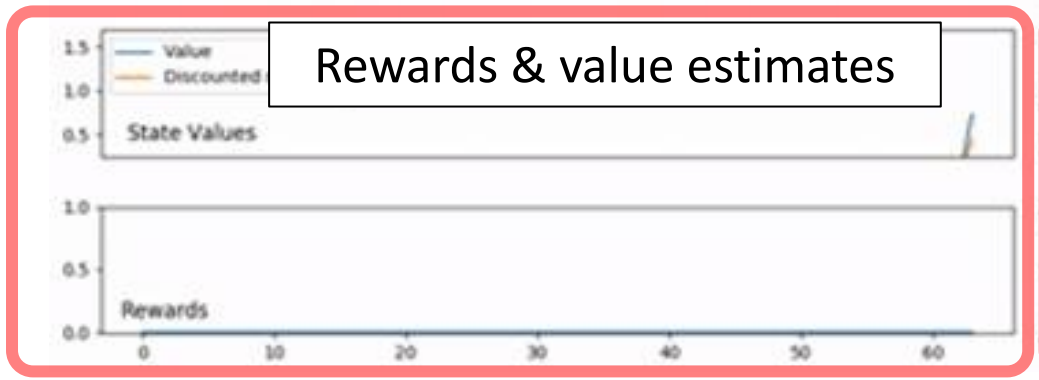
Mirowski et al. ICLR 2017, Jaderburg et al. ICLR 2017, Mirowski et al . ICLR 2018

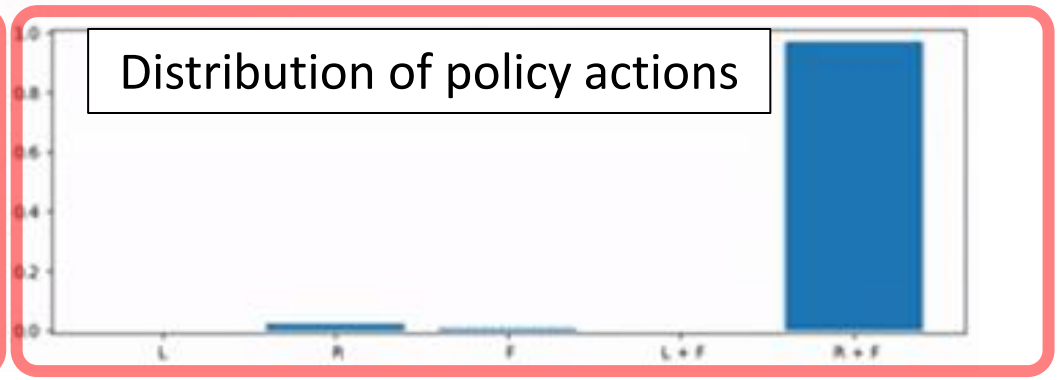
Example task: K-item scenario

Top down map

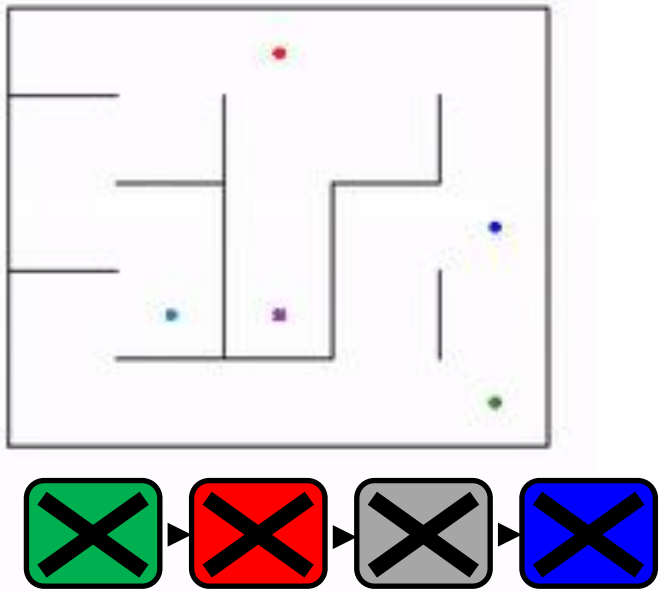


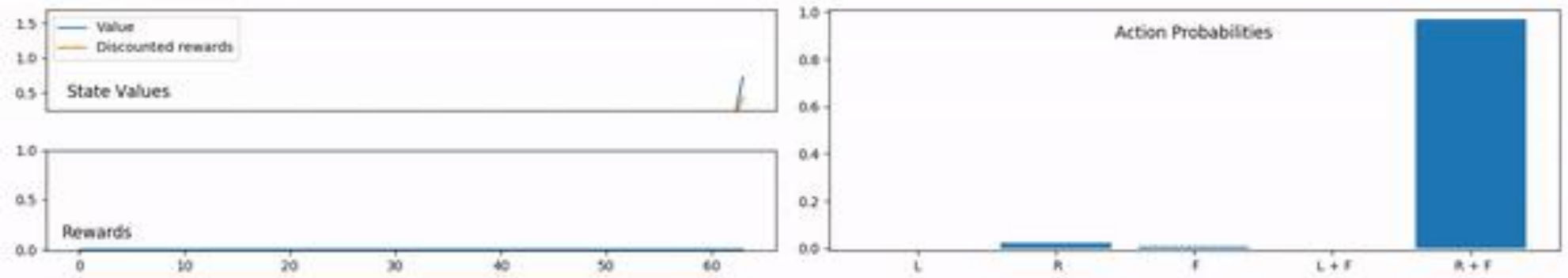
Items





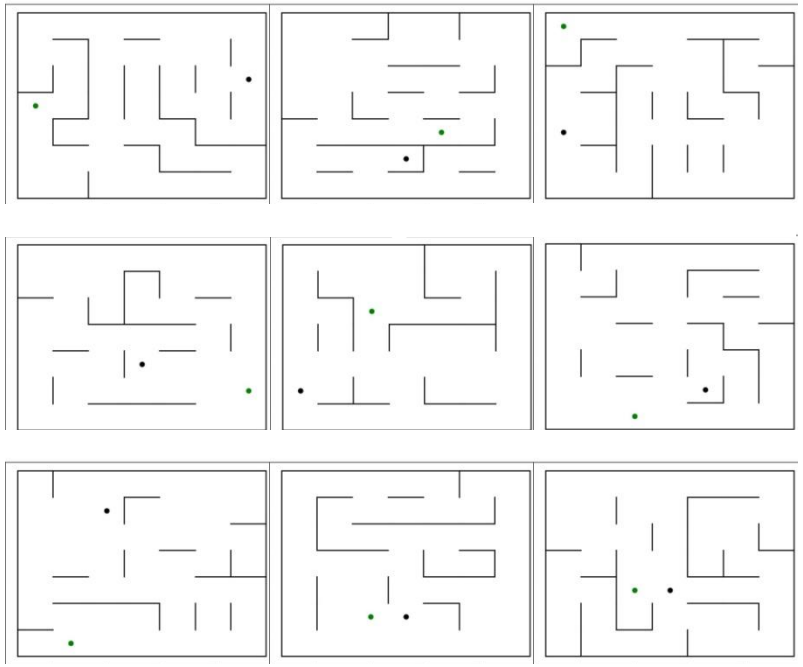
Example task: K-item scenario



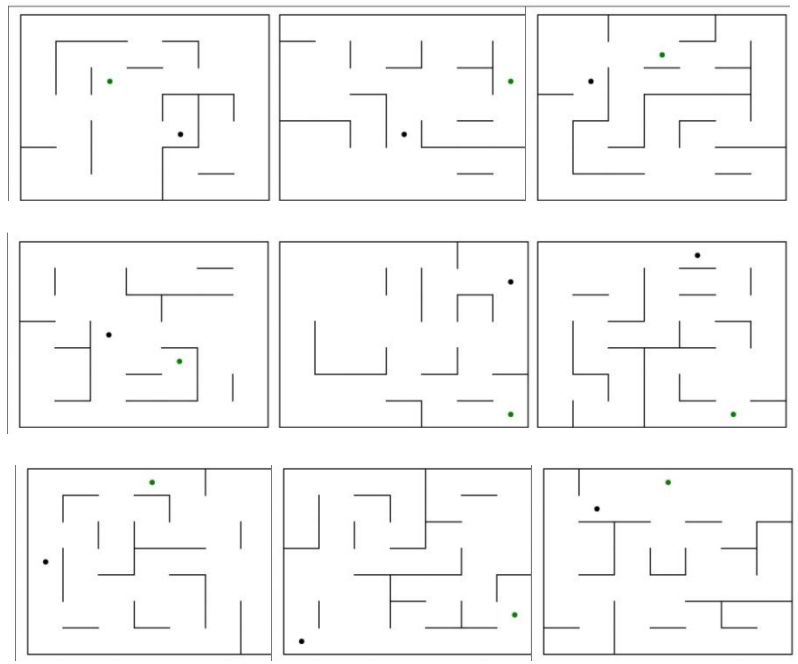


Generalization

N - Training configurations

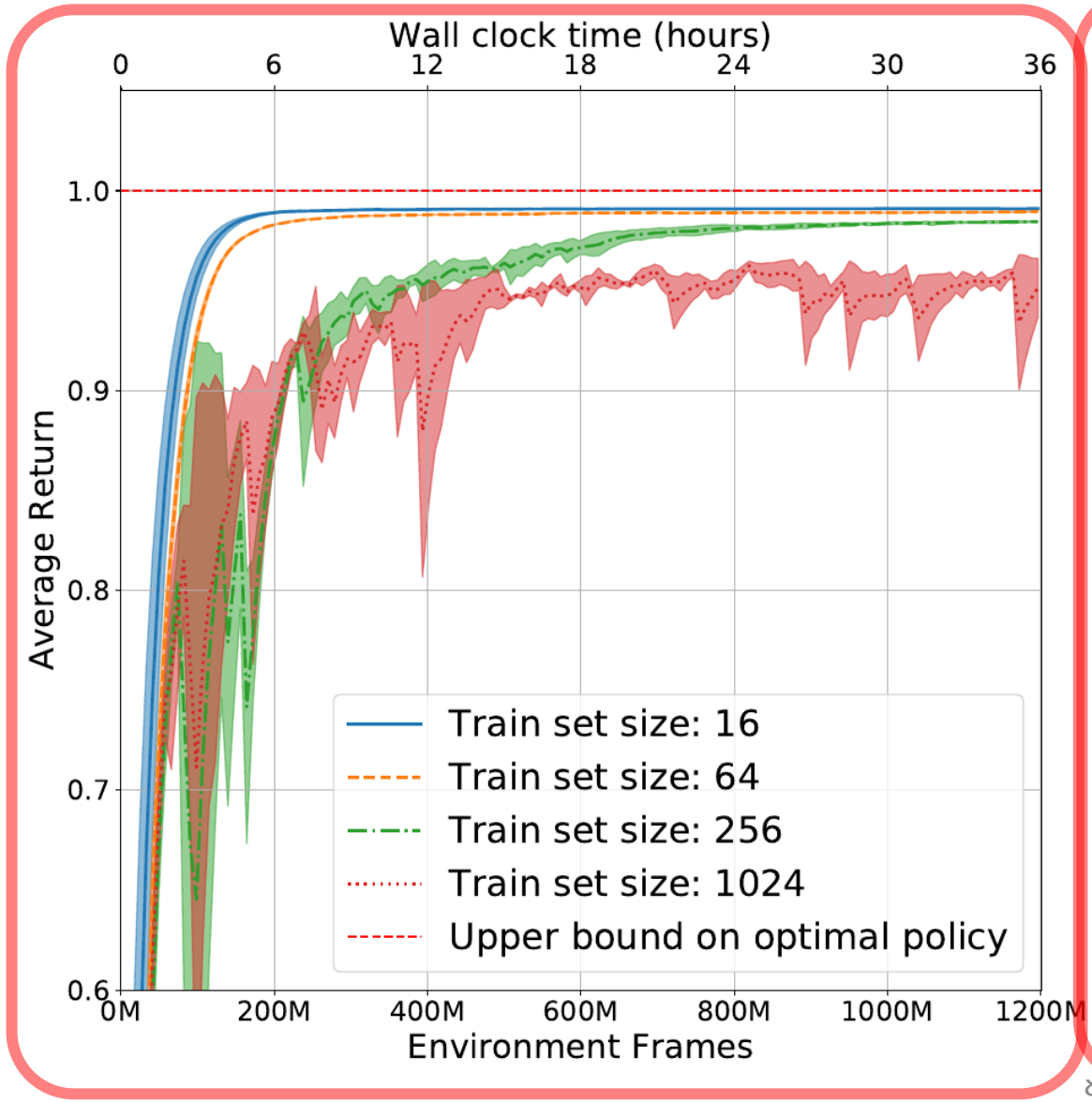


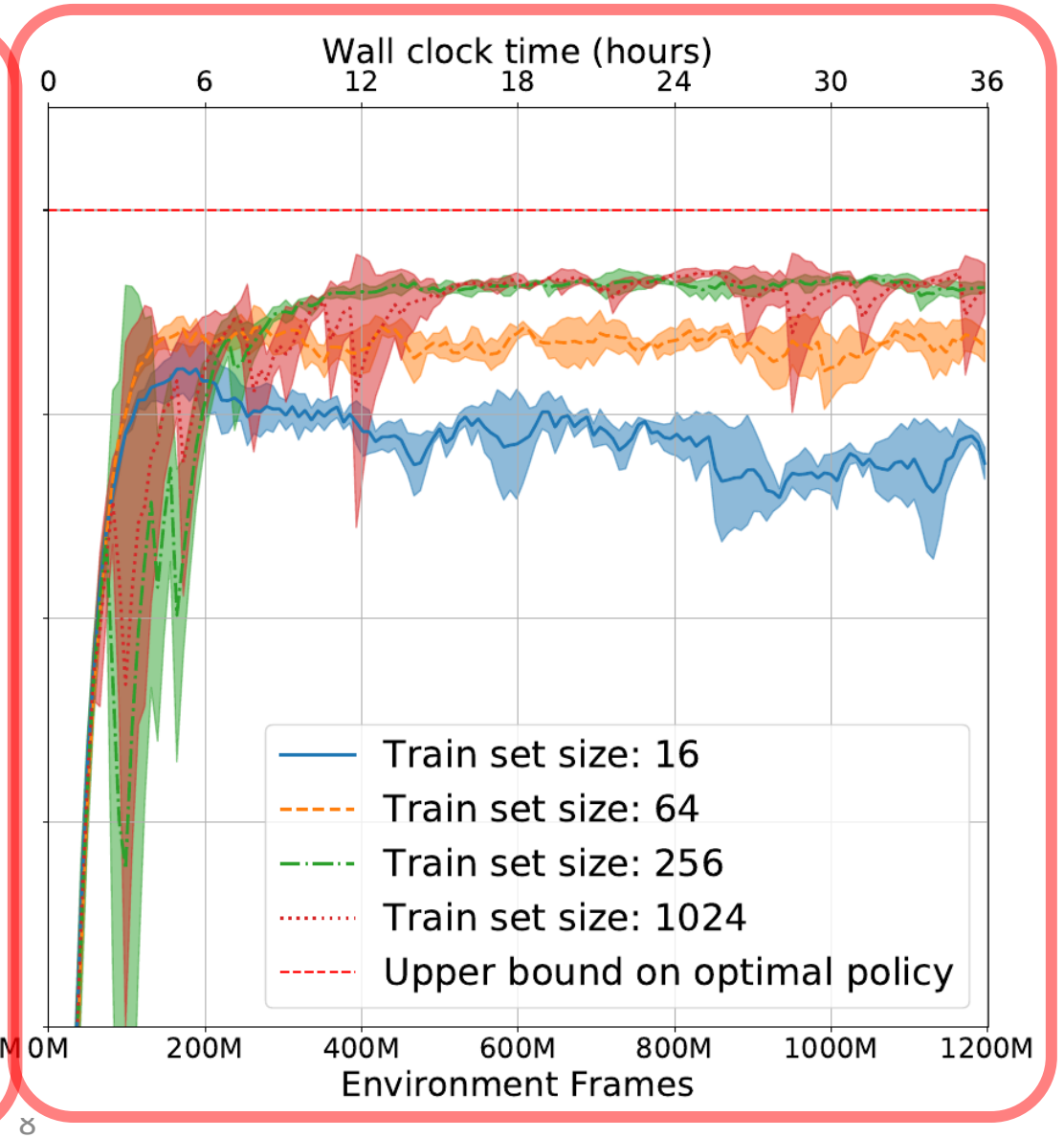
K - Testing configurations



How many training configurations are required to generalize to unseen test configurations?

Generalization results





Results

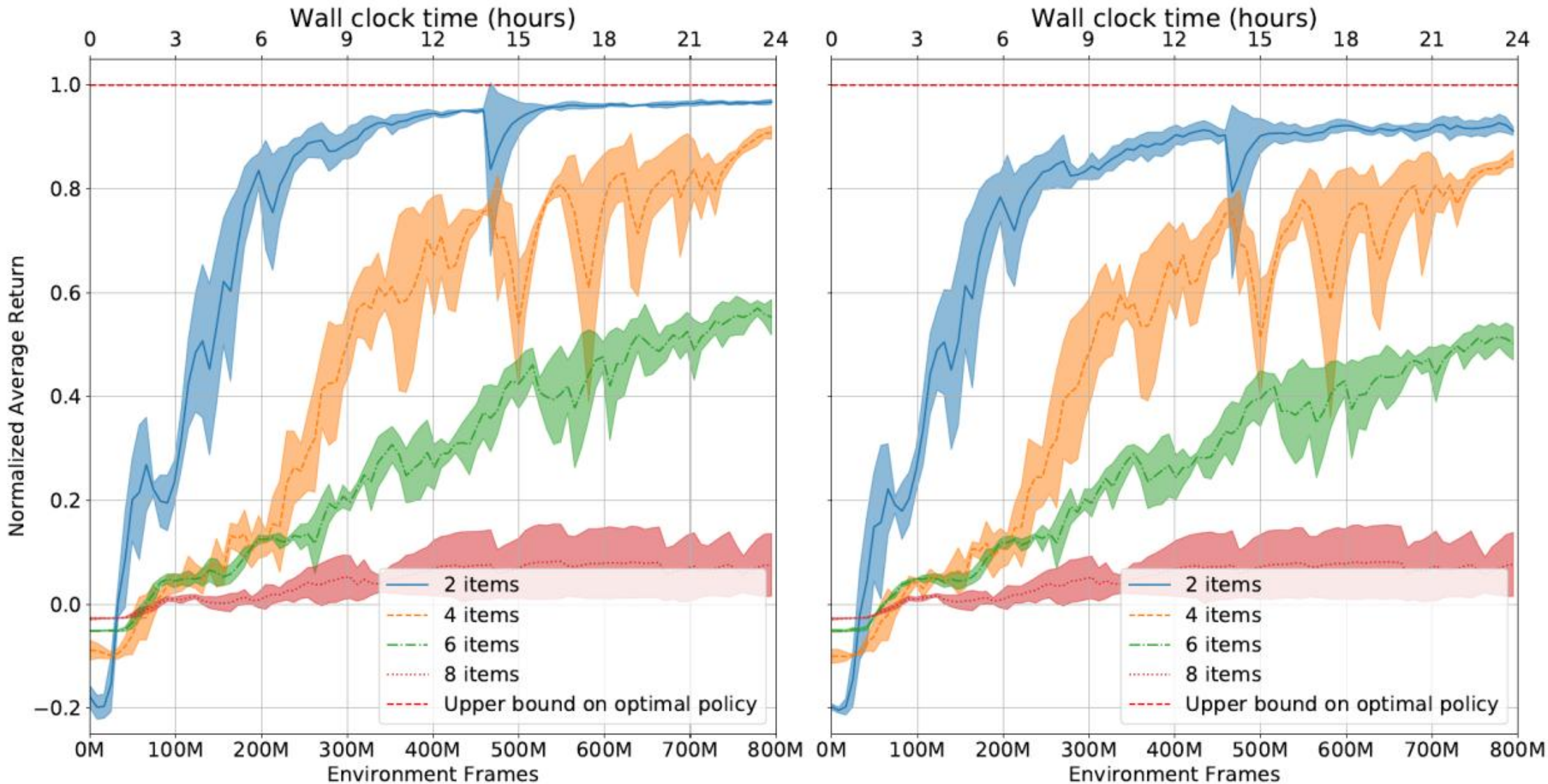
Labyrinth		
Size	**Train**	**Test**
7	99.8 ± 0.2	99.5 ± 0.2
9	98.1 ± 0.8	97.4 ± 1.3
11	95.2 ± 0.7	90.1 ± 1.0
13	84.1 ± 1.8	79.8 ± 1.9

Ordered K-item		
# items	**Train**	**Test**
2	0.968 ± 0.004	0.932 ± 0.012
4	0.910 ± 0.012	0.861 ± 0.015
6	0.577 ± 0.024	0.522 ± 0.023
8	0.084 ± 0.071	0.083 ± 0.070

Find Return		
Size	**Train**	**Test**
7	98.7 ± 0.3	95.1 ± 0.2
9	87.0 ± 1.9	81.6 ± 1.8
11	70.9 ± 1.6	64.0 ± 1.6
13	57.7 ± 0.8	52.9 ± 3.7

Two colors		
Complexity	**Train**	**Test**
1	1903 ± 27	1941 ± 4
3	1789 ± 24	1781 ± 24
5	1436 ± 158	1432 ± 161
7	1159 ± 126	1128 ± 140

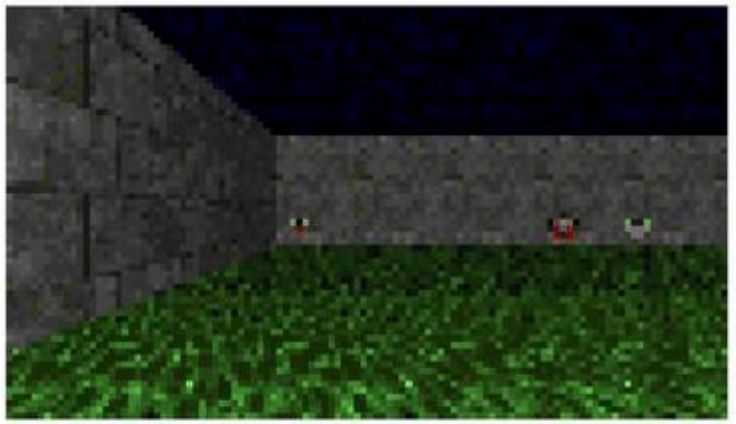
Results: K-item



This work

- We provide challenging scenarios for Deep Reinforcement Learning agents in partially observable 3D environments.

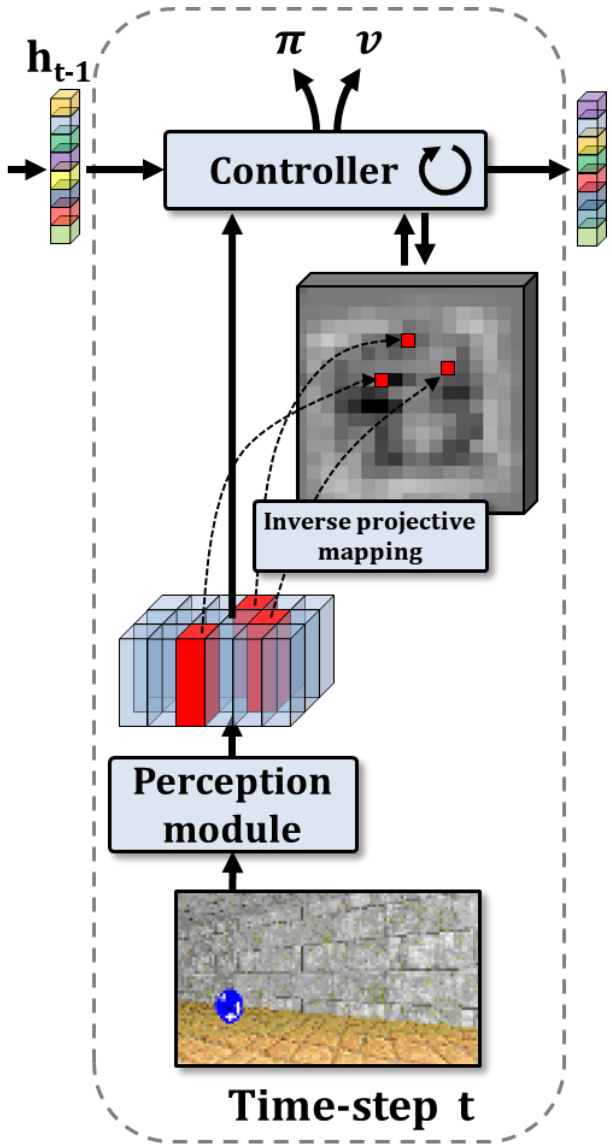
- This work:
  - Extends the ViZDoom simulator with 4 new scenarios
  - Trains and evaluates a recurrent Deep RL agent to identify its limitations and areas for improvement
  - Evaluates the generalization performance in unseen environment configurations.
- Our scenarios:
  - Act as a proxy to real world environments
  - Test reasoning, memory, generalization and planning in complex 3D environments.
  - Provide a standard benchmark for Deep RL algorithms that can be trained in under 24 hours



Extensions

EgoMap: Projective mapping and structured egocentric memory for Deep RL. ECML-PKDD 2020

- Extends recurrent baselines with structured memory
- Incorporates differentiable inverse projective mapping
- Outperforms recent metric map methods



Deep Reinforcement Learning on a Budget
3D Control and reasoning without a supercomputer

ICPR 2020

Webpage: https://edbeeching.github.io/papers/3d_control_deep_rl

See you in the poster session!




Edward Beeching

Jilles Dibangoye

Olivier Simonin

Christian Wolf