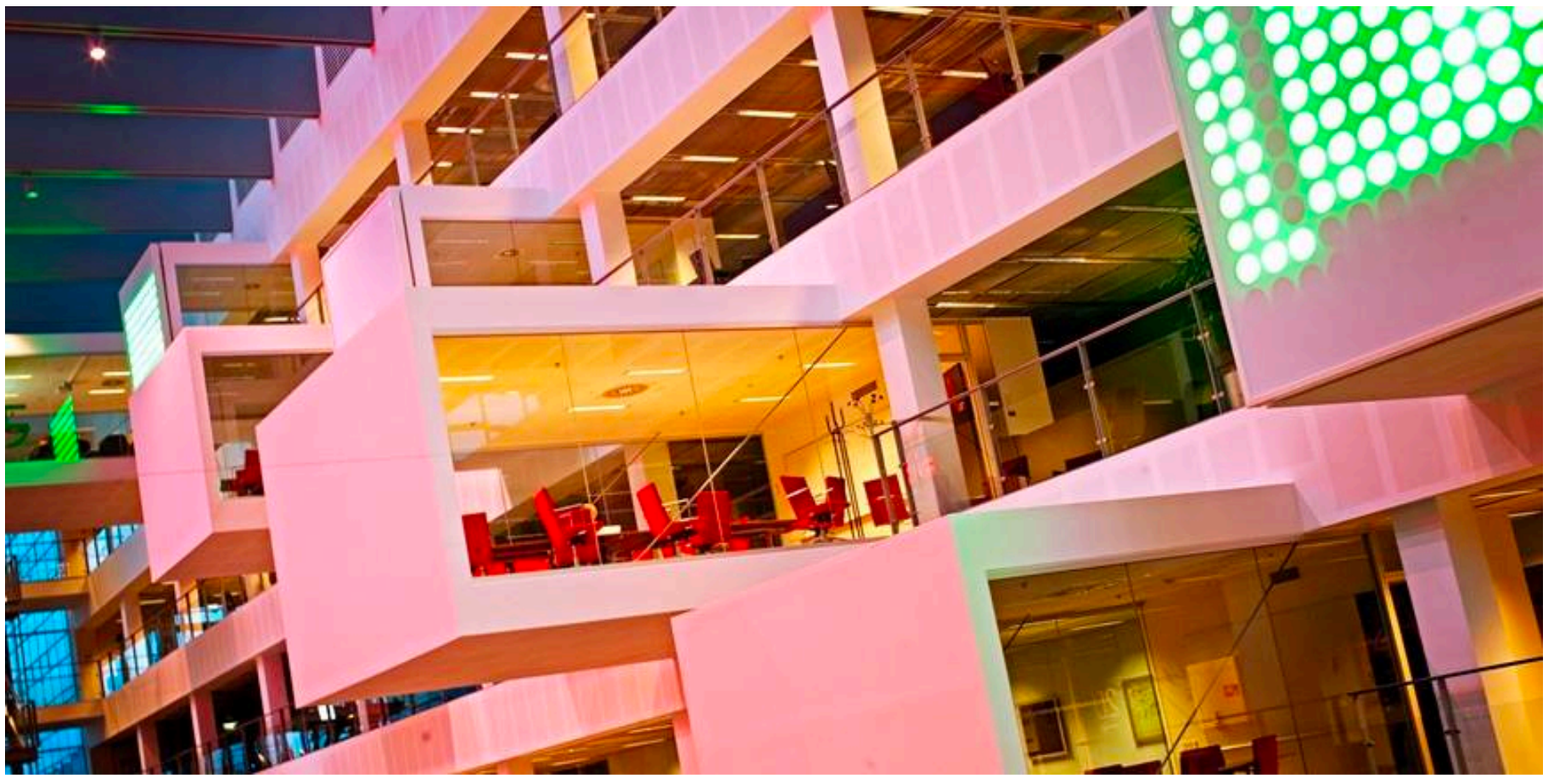
Adaptive AI for Games

Niels Justesen

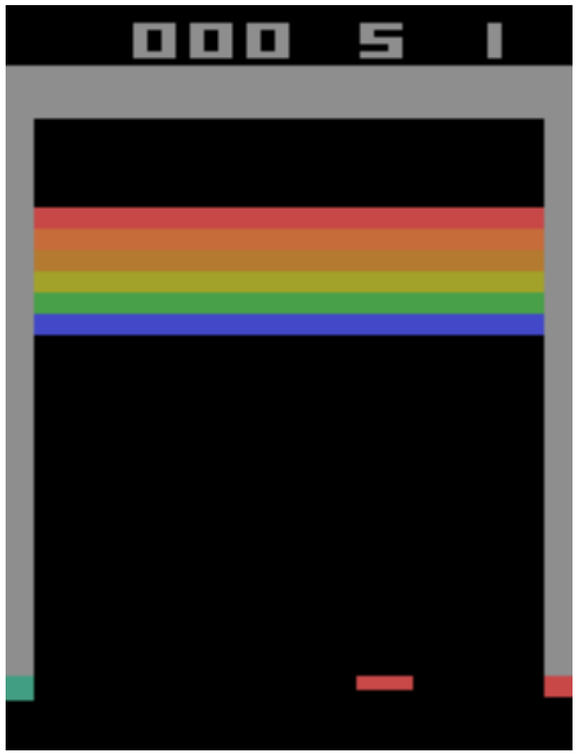
Visiting PhD student from the IT University of Copenhagen

Robotics, Evolution, and Art Lab (REAL)

Center for Computer Games Research



How can algorithms (learn to) play advanced video games?



?

Adaptive Behaviours in Games

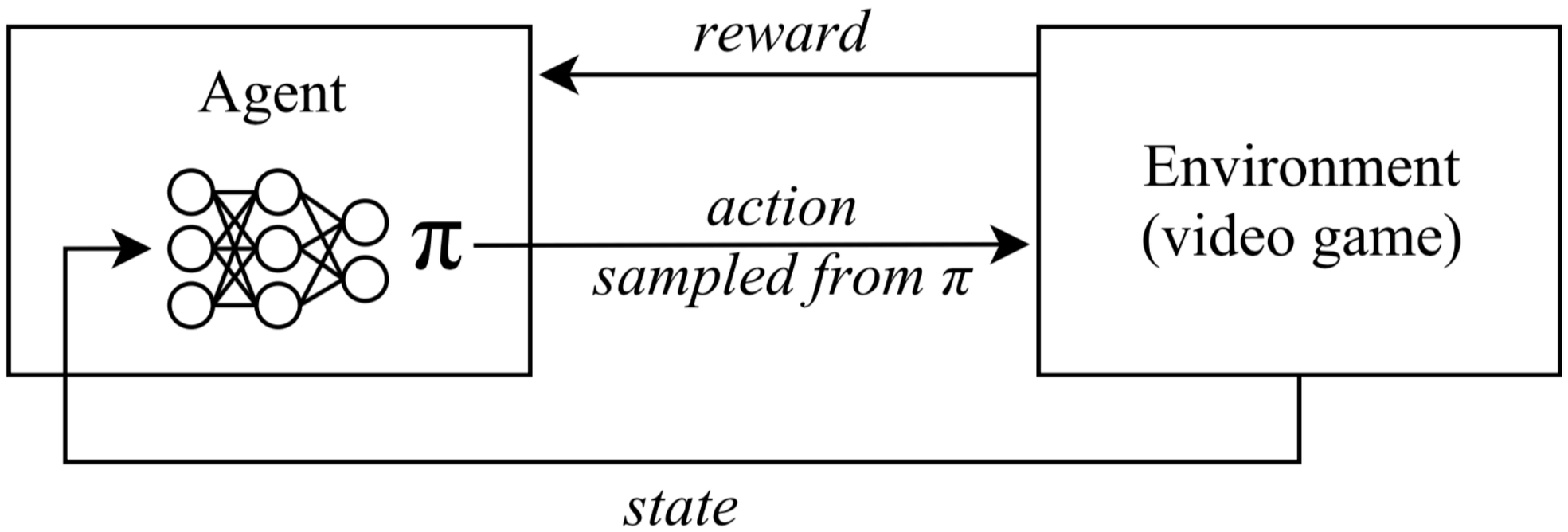
Hierarchy of decisions in Games:

- High-level decisions (strategy/macro)
- Low-level decisions (control/micro)

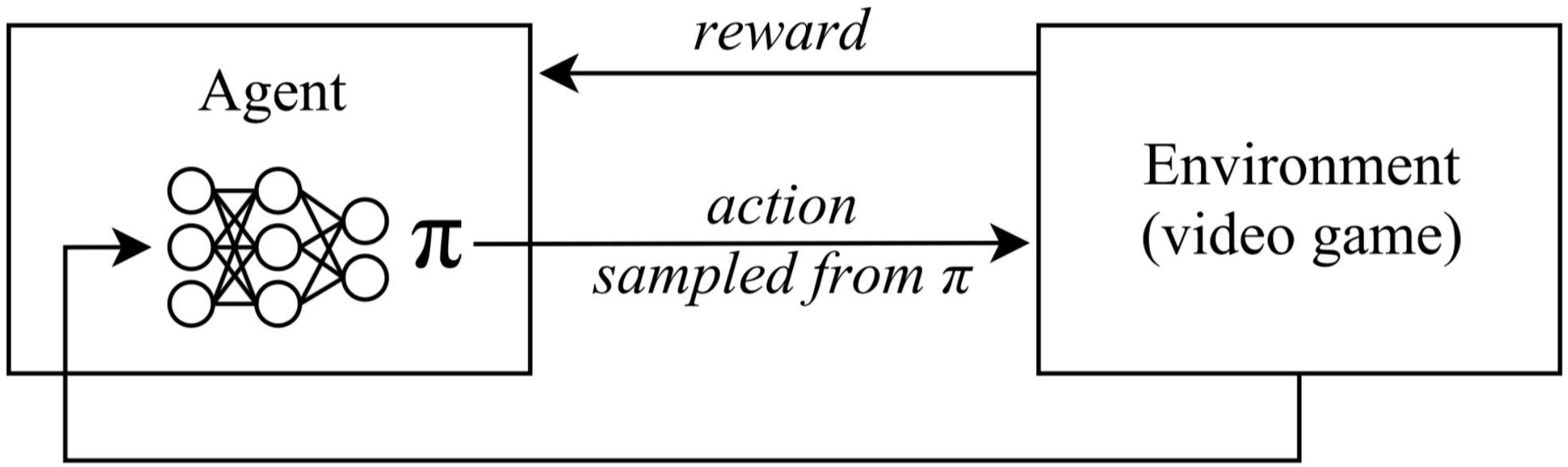
Types of adaptivity in Games:

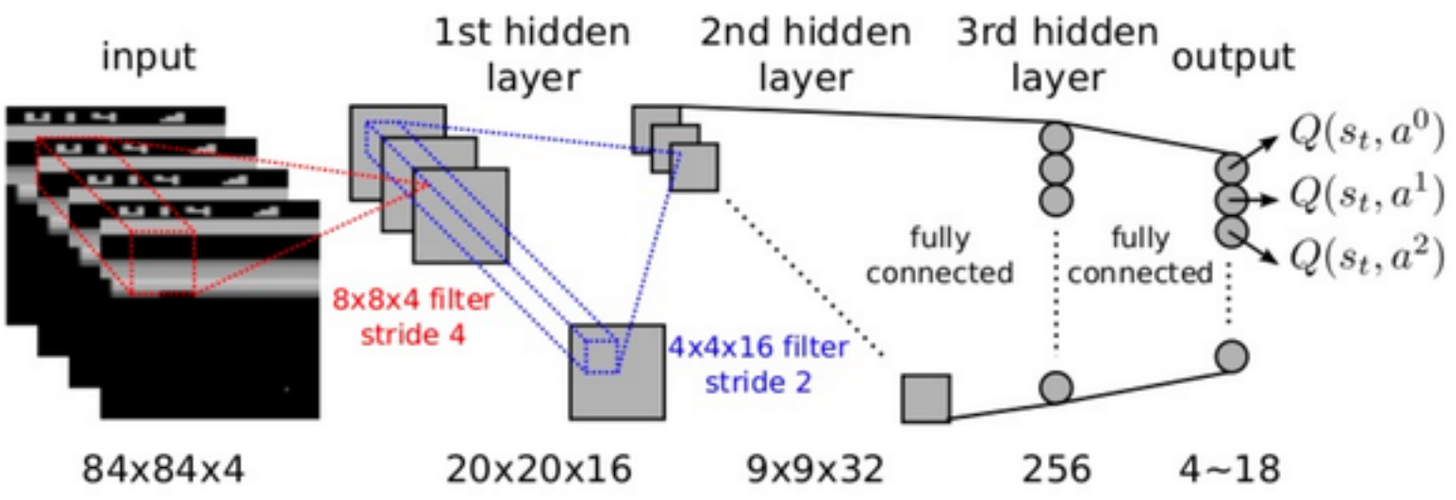
- **Inter-game** adaptive behaviours
- **Intra-game** adaptive behaviours
- **Balanced** behaviours
- **Non-adaptive** behaviours

Deep Reinforcement Learning



Deep Reinforcement Learning

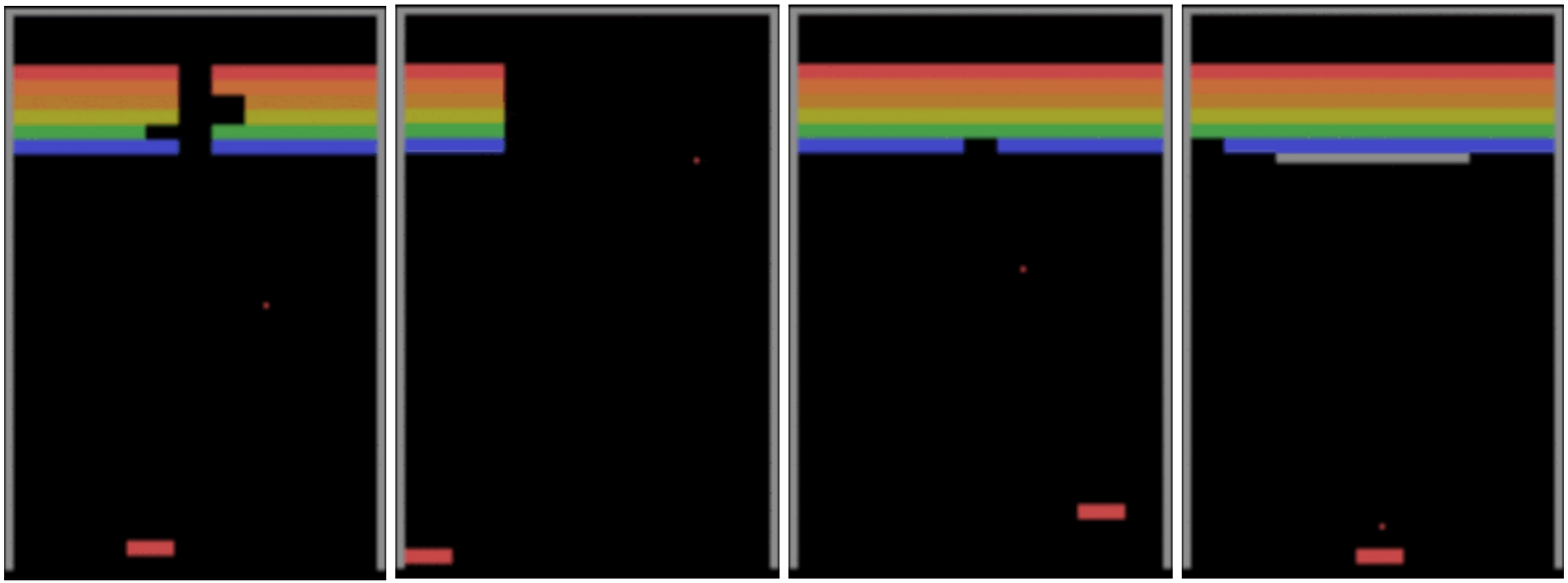




Non-adaptive Behaviours

Training level

Test levels



Breakout
Normal

Breakout
Random Target

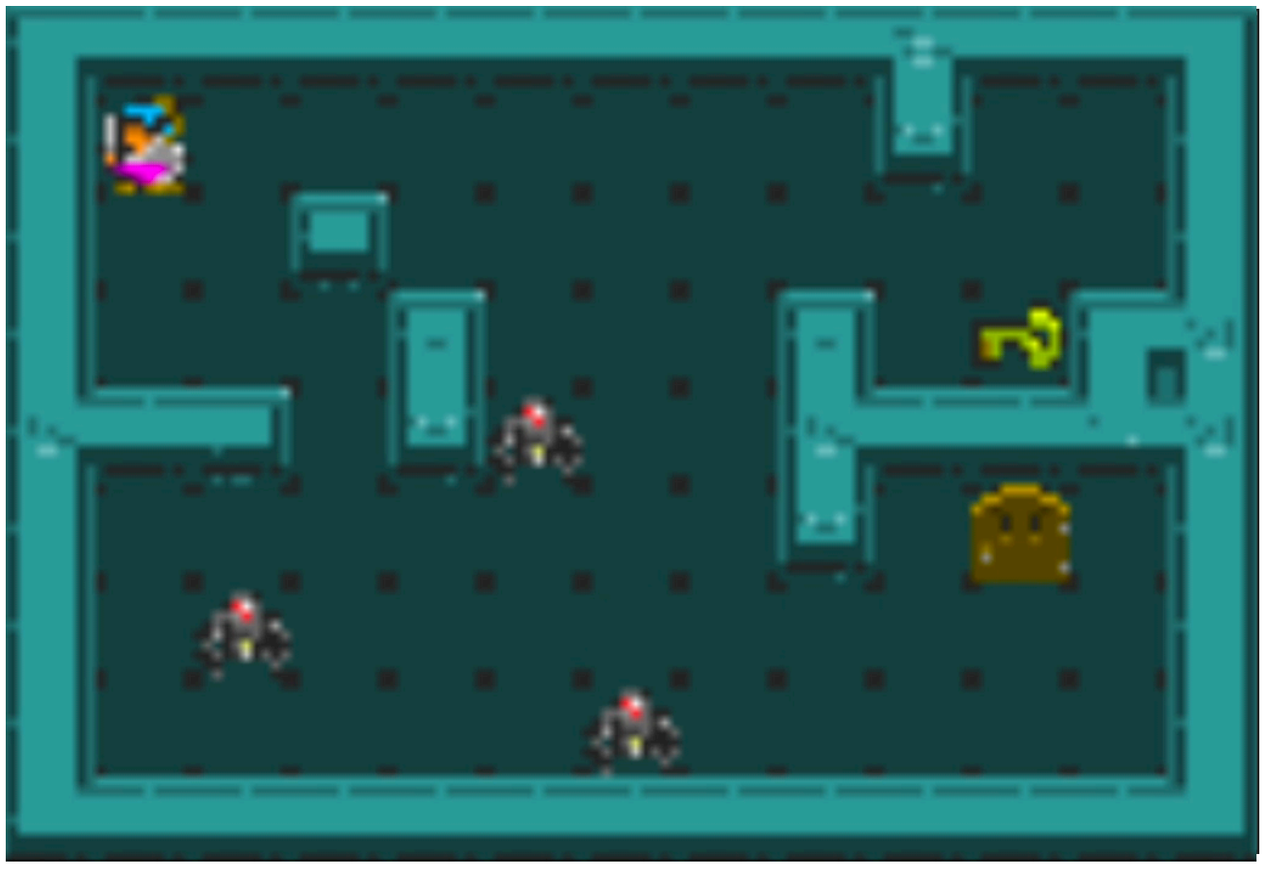
Breakout
Higher paddle

Breakout
Wall

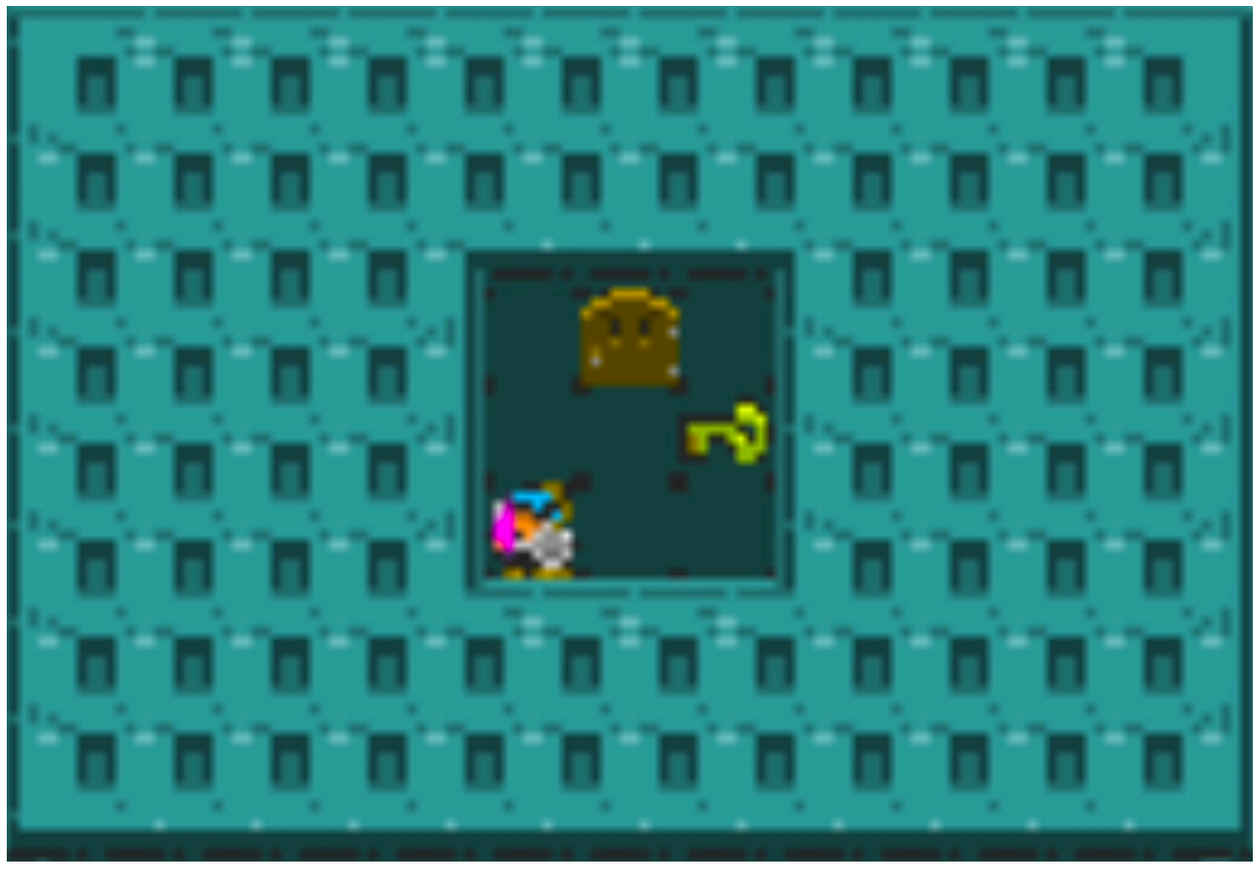
Kansky, Ken, et al. "Schema networks: Zero-shot transfer with a generative causal model of intuitive physics." *arXiv preprint arXiv:1706.04317* (2017).

Non-adaptive Behaviours

Training level

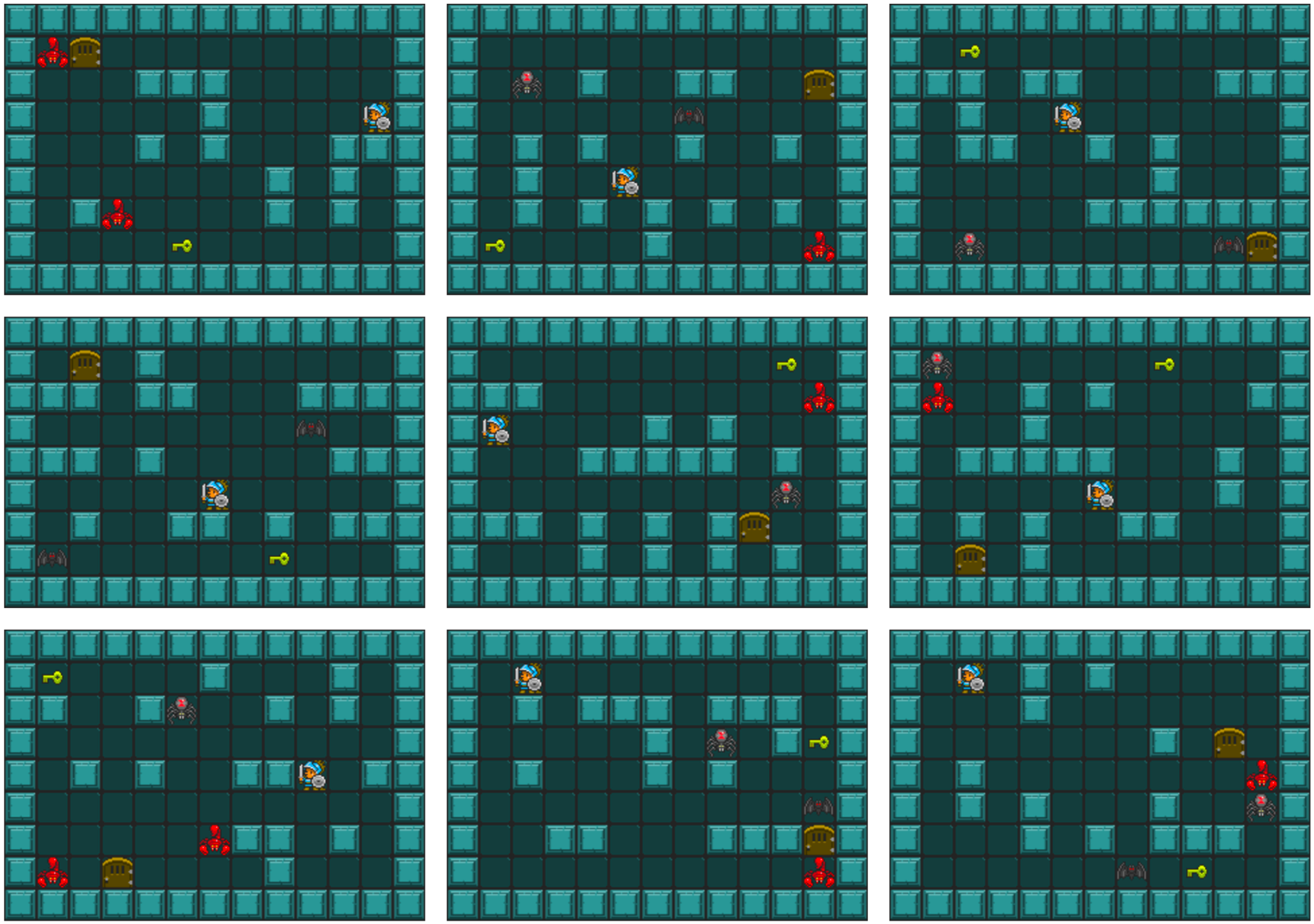


Test levels



Justesen, Niels, et al. "Illuminating Generalization in Deep Reinforcement Learning through Procedural Level Generation." *arXiv preprint arXiv:1806.10729* (2018).

Procedural Level Generation

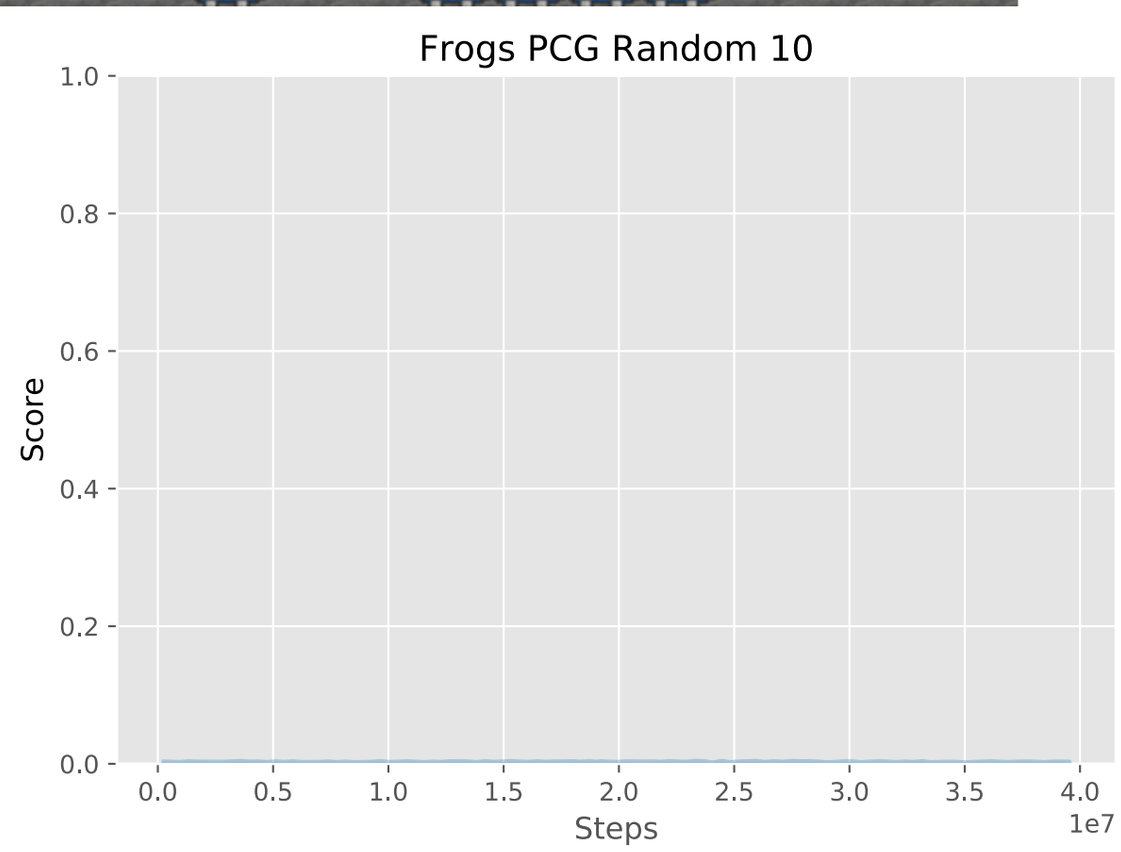


Justesen, Niels, et al. "Illuminating Generalization in Deep Reinforcement Learning through Procedural Level Generation." *arXiv preprint arXiv:1806.10729* (2018).

Procedural Level Generation + RL

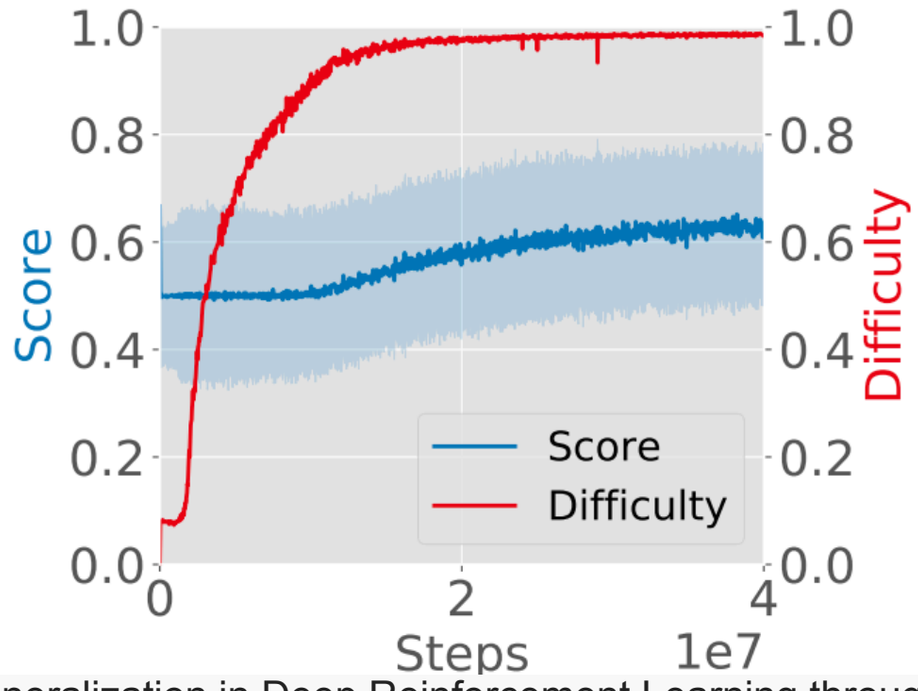
Justesen, Niels, et al. "Illuminating Generalization in Deep Reinforcement Learning through Procedural Level Generation." *arXiv preprint arXiv:1806.10729* (2018).

PCG in Games with Sparse Rewards



Justesen, Niels, et al. "Illuminating Generalization in Deep Reinforcement Learning through Procedural Level Generation." *arXiv preprint arXiv:1806.10729* (2018).

PCG + Curriculum Learning + RL

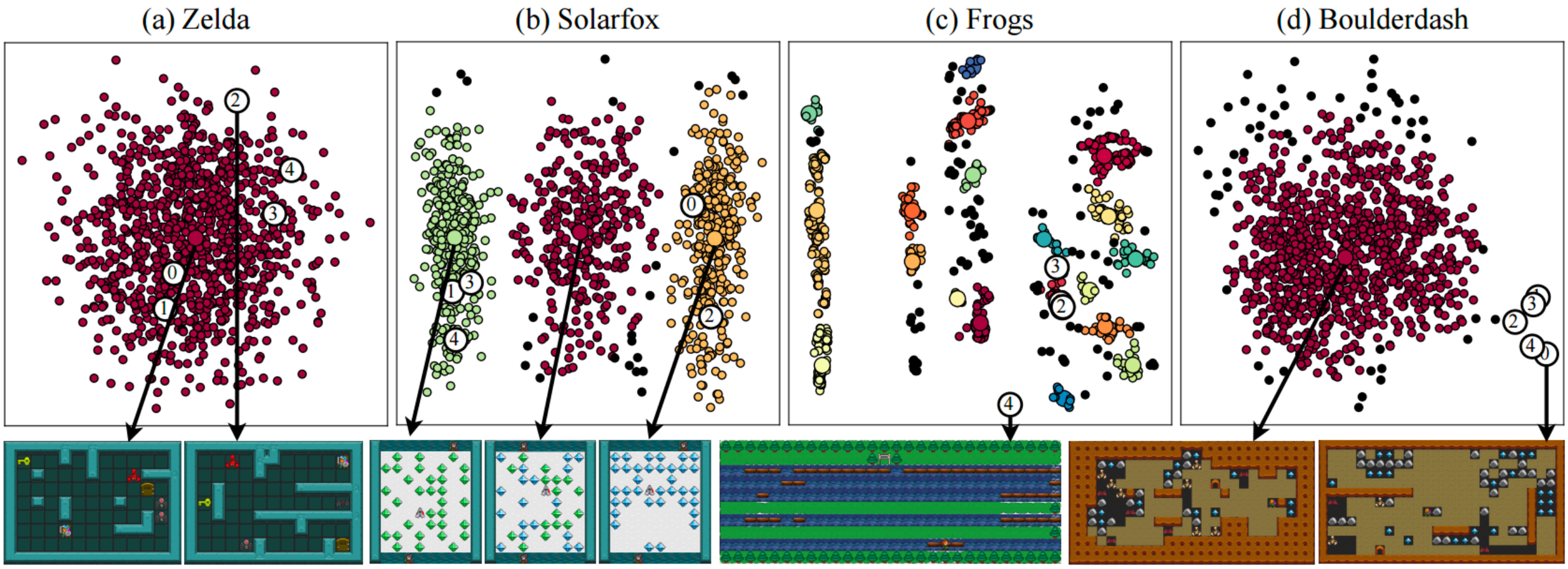


Justesen, Niels, et al. "Illuminating Generalization in Deep Reinforcement Learning through Procedural Level Generation." *arXiv preprint arXiv:1806.10729* (2018).

Human-designed Test Level

Justesen, Niels, et al. "Illuminating Generalization in Deep Reinforcement Learning through Procedural Level Generation." *arXiv preprint ar*
1806.10729 (2018).

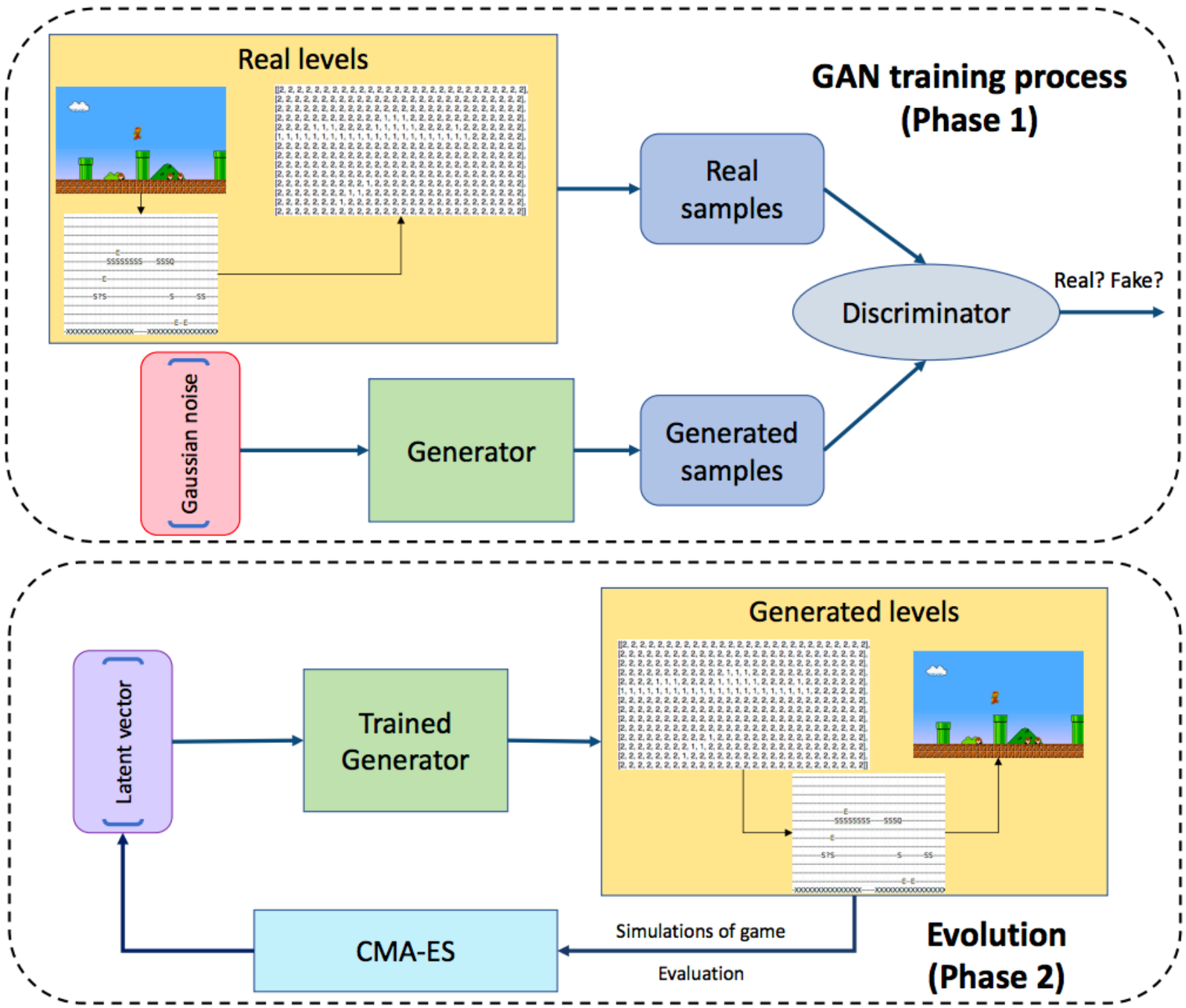
Level Distributions



1. **Principal Component Analysis (PCA)**
2. **Density-Based Spatial Clustering of Applications with Noise (DBSCAN)**

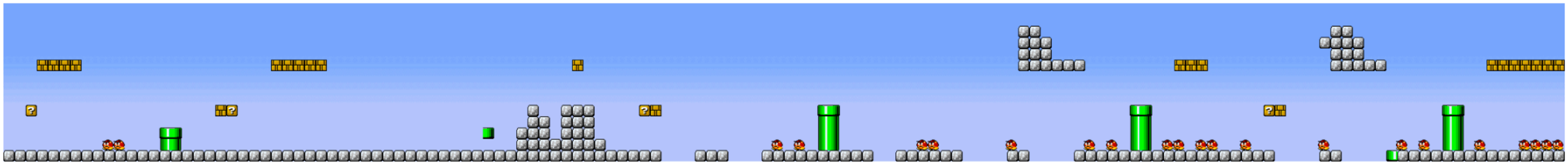
Justesen, Niels, et al. "Illuminating Generalization in Deep Reinforcement Learning through Procedural Level Generation." *arXiv preprint ar*
1806.10729 (2018).

Learning to Design Levels from Examples



Volz, V.; Schrum, J.; Liu, J.; Lucas, S. M.; Smith, A.; and Risi, S. 2018. Evolving mario levels in the latent space of a deep convolutional generative adversarial network. arXiv preprint arXiv:1805.00728.

Learning to Design Levels from Examples



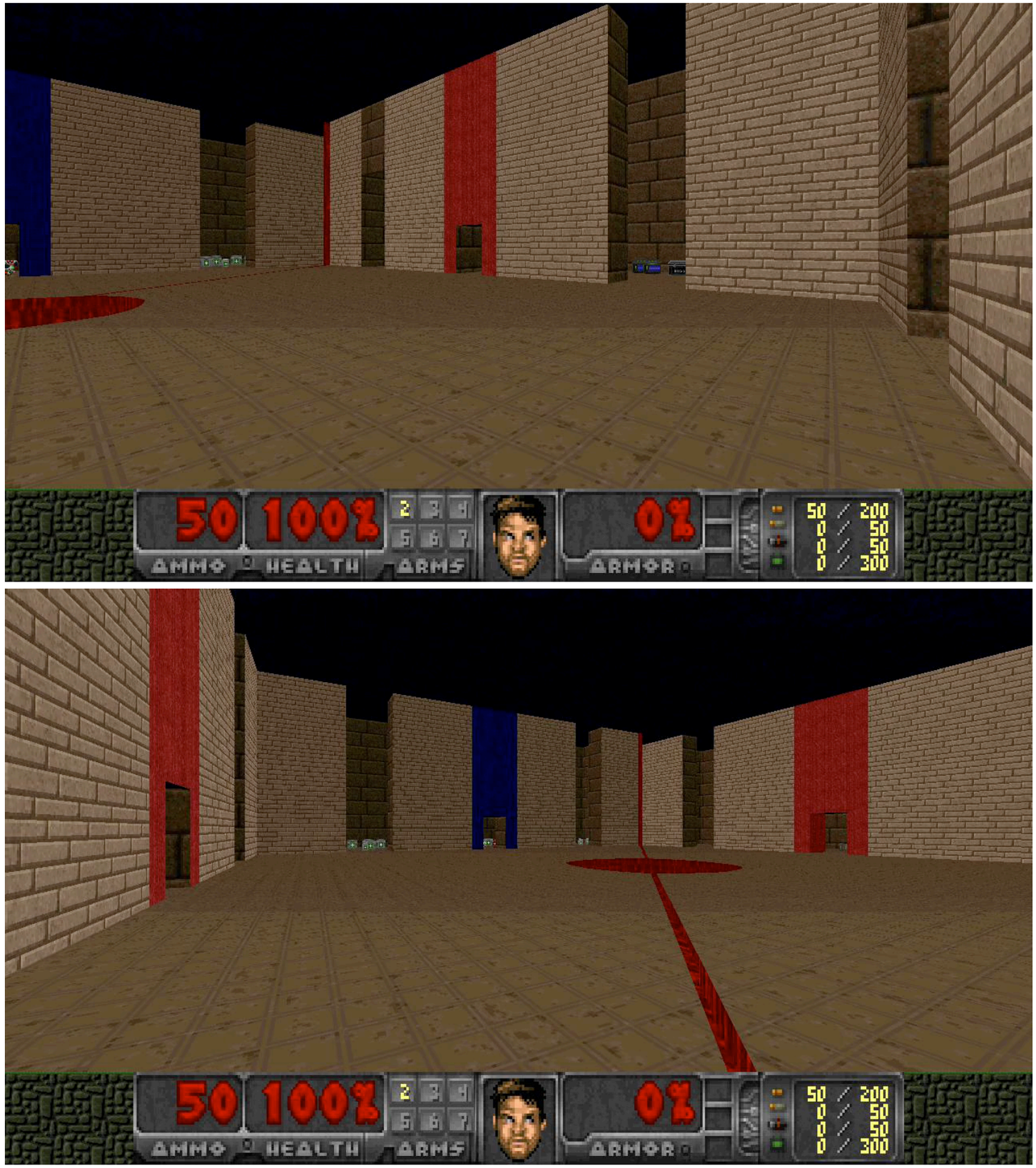
Volz, V.; Schrum, J.; Liu, J.; Lucas, S. M.; Smith, A.; and Risi, S. 2018. Evolving mario levels in the latent space of a deep convolutional generative adversarial network. arXiv preprint arXiv:1805.00728.

Intra-game Adaptive Build-orders in StarCraft

Successful tricks when applying GA's for online planning

- Populations restarts
- Keep *k* best solutions (freezing)
- EA-shift / continual online evolution

Non-adaptive “Overfitted” Behaviour in Doom



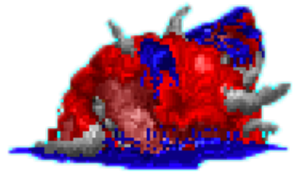
Justesen, Niels, and Sebastian Risi. "Automated Curriculum Learning by Rewarding Temporally Rare Events." *Computational Intelligence and Games (CIG), IEEE* (2018).

Manual Reward Shaping

1

10

5



100

5

10

15

25

50

Manual Reward Shaping

1 **10** **5** **100**

5 **10** **15** **25** **50**

Rewarding temporally rare events

				μ	R
	6	10	8	8	0.125
	1	0	1	0.67	1.492
	.	.	.	.	.
	.	.	.	.	.
	.	.	.	.	.
	0	1	0	0.33	3

Event buffer

$$R_t(\epsilon) = \frac{1}{\mu_t(\epsilon)}$$

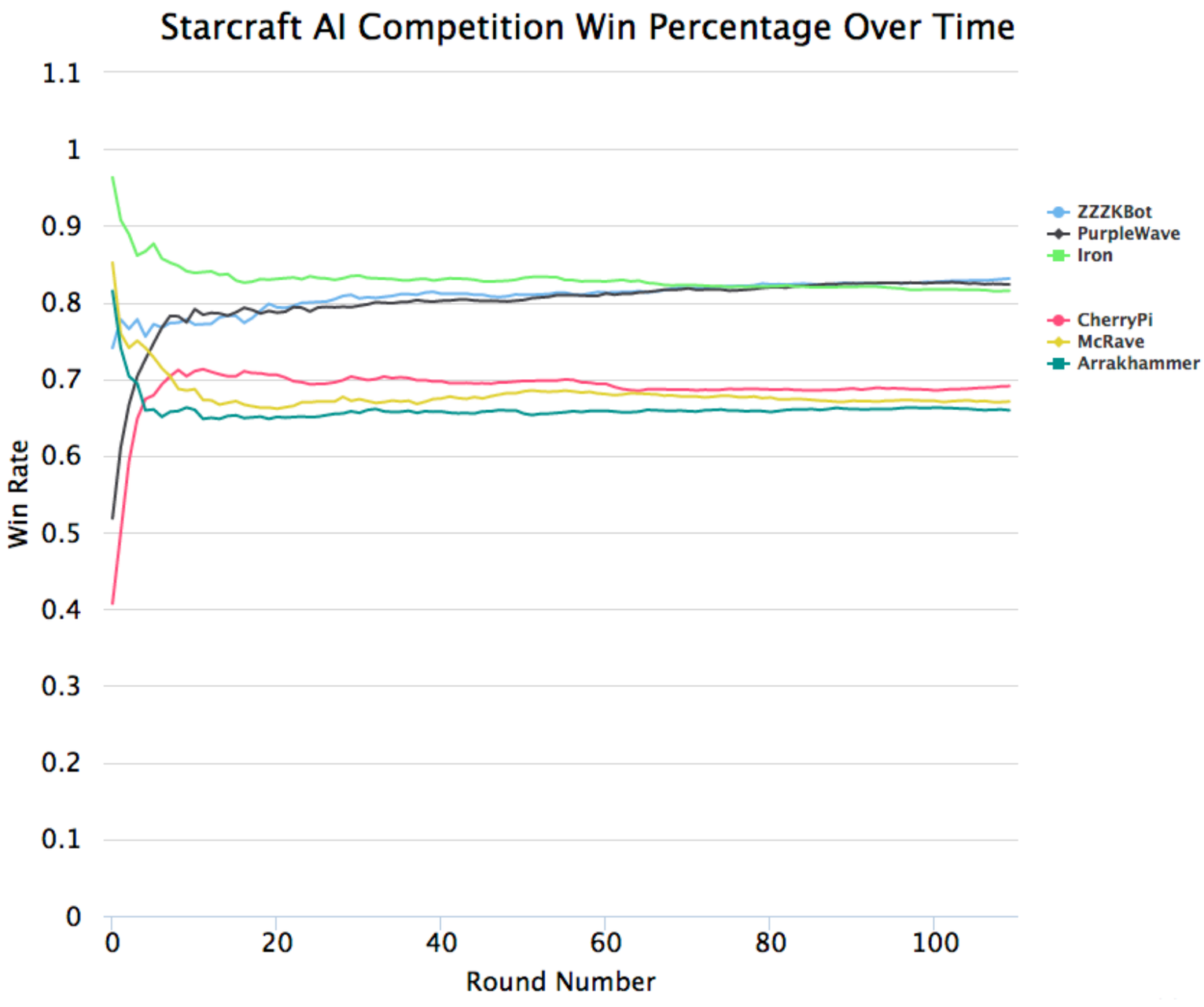
Balanced Behaviours

Rewarding RL agents by temporally rare events results in
a **balanced behaviour** which shows **improved generalisation.**

Justesen, Niels, and Sebastian Risi. "Automated Curriculum Learning by Rewarding Temporally Rare Events." *Computational Intelligence and Games (CIG), IEEE* (2018).

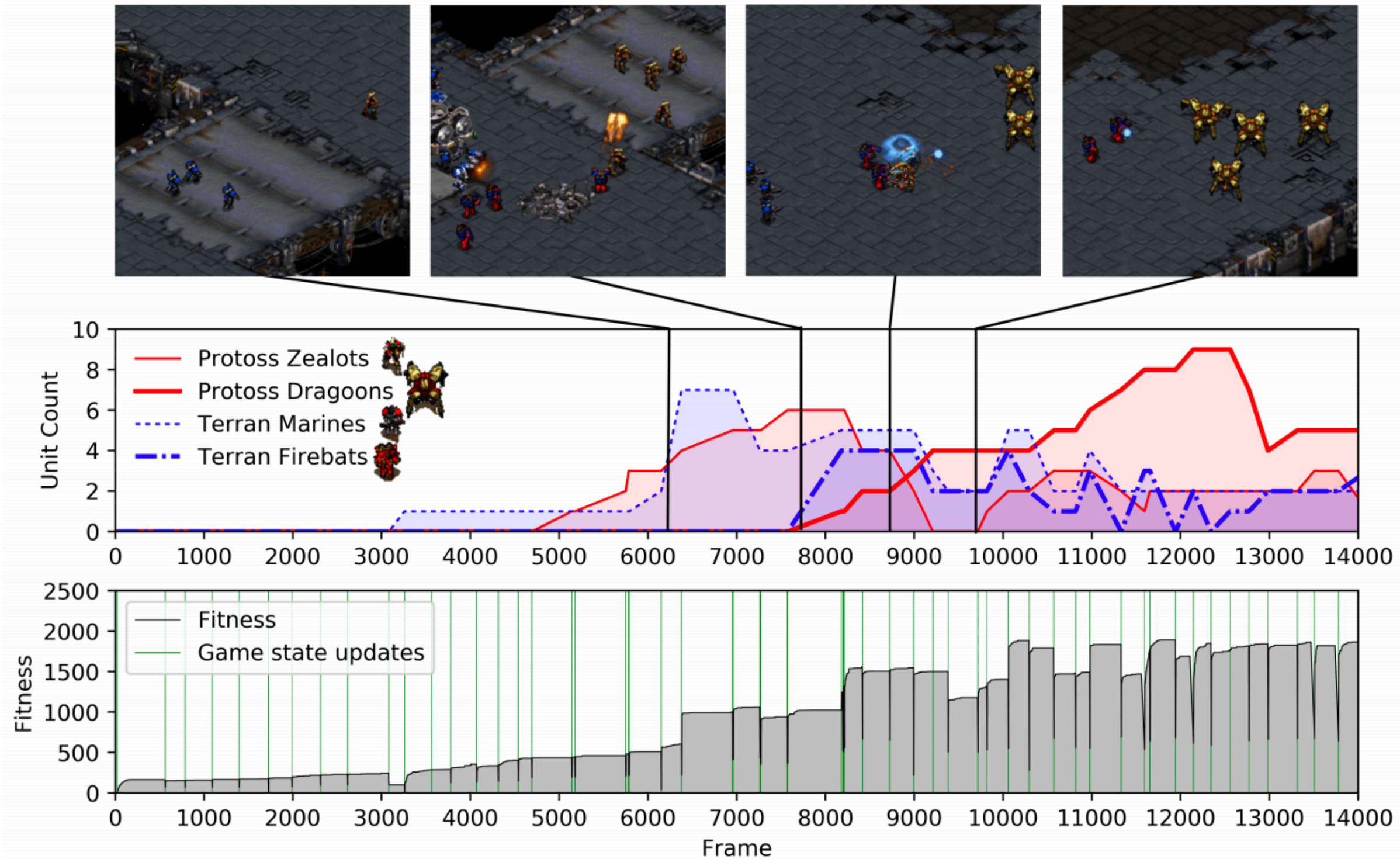
Inter-game Adaptive Behaviour

When we play the same opponents multiple times,
we can treat strategy selection as a multi-armed bandit problem.



http://www.cs.mun.ca/~dchurchill/starcraftaicomp/2017/

Intra-game Adaptive Build-orders in StarCraft



TStarBots: Defeating the Cheating Level Builtin AI in StarCraft II in the Full Game

Peng Sun*[1], Xinghai Sun*[1], Lei Han*[1], Jiechao Xiong*[1], Qing Wang[1], Bo Li[1], Yang Zheng[1], Ji Liu[1,2], Yongsheng Liu[1], Han Liu[1,3], Tong Zhang[1]

[1]Tencent AI Lab [2]University of Rochester [3]Northwestern University

September 20, 2018

Non-adaptive Behaviour in StarCraft

*“It is worthy noting that although TStarBot1 can successfully learn and acquire strategies to defeat all the builtin AIs, it lacks **diversity** in order to consistently beat human players.”*

“In the aforementioned test with human players, TStarBot1 will be unable to win once the human player starts to know TStarBot1’s preference for Zergling Rush.”

Sun, Peng, et al. "TStarBots: Defeating the Cheating Level Builtin AI in StarCraft II in the Full Game." *arXiv preprint arXiv:1809.07193* (2018).

Non-adaptive Behaviour in StarCraft

Sun, Peng, et al. "TStarBots: Defeating the Cheating Level Builtin AI in StarCraft II in the Full Game." *arXiv preprint arXiv:1809.07193* (2018).