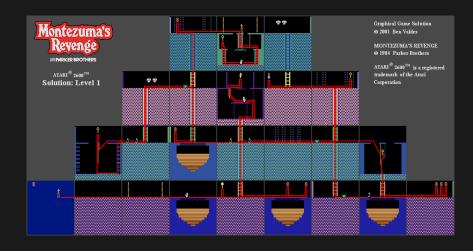


Montezuma's Revenge



Montezuma's Revenge



- very sparse rewards hundrets of steps
- huge state space
- hard exploration
- needs returns back

highlighted score

name

vear

https://papers with code. com/sota/atari-games-on-atari-2600-montez umas revenge

| J | | |
|------|---|--------|
| 2015 | Deep Reinforcement Learning with Double Q-learning | 0 |
| 2021 | MuZero | 2500 |
| 2018 | Count-Based Exploration with Neural Density Models ¹ | 3705 |
| 2019 | Exploration by Random Network Distillation ² | 8152 |
| 2021 | GoExplore* ³ | 43 000 |

* : requires environment state saving/loading

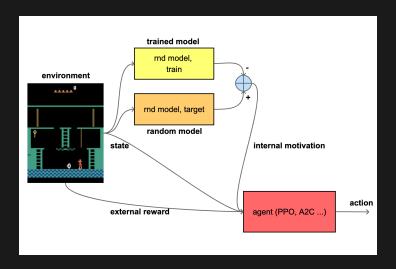
score

 $^{^{1}\}mathsf{https://arxiv.org/abs/1703.01310}$

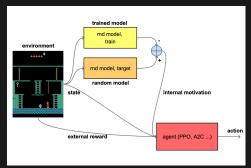
²https://arxiv.org/abs/1810.12894

³https://arxiv.org/abs/2004.12919

random network distillation



random network distillation



- neural network works as novelty detector
- model learns to imitate random (target) model
- less visited states produce bigger motivation signal
- orthogonal weights initialisation $(g = 2^{0.5})$ for strong signal
- lot of fully connected layers to avoid generalisation

Q&A

