

1. Informacje ogólne

Tytuł projektu: Czy na MuZero można odpalić DOOM'a?

Podstawa projektu: Mastering Atari, Go, Chess and Shogi by Planning with a Learned Model

<https://arxiv.org/pdf/1911.08265>

Julian Schrittwieser, Ioannis Antonoglou, Thomas Hubert, Karen Simonyan, Laurent Sifre, Simon Schmitt, Arthur Guez, Edward Lockhart, Demis Hassabis, Thore Graepel, Timothy Lillicrap, David Silver

Repozytorium:

https://github.com/MarekBorzyszkowski/MuZero_redone

2. Cel projektu

Próba powtórzenia eksperymentu z publikacji i rozszerzenie eksperymentu o nowe środowisko DOOM.

3. Przegląd artykułu

Model/Algorytm bazowy: MuZero

Streszczenie artykułu: Przedstawienie algorytmu MuZero i porównanie jego osiągnięć względem istniejących rozwiązań.

Kluczowe wnioski: MuZero radzi sobie zarówno z grami wymagającymi planowania (takimi jak szachy, czy Go) oraz posiadającymi skomplikowany interfejs graficzny. MuZero łączy planowanie z brakiem odtworzonego w pełni środowiska, posiadając tylko dań istotne informacje w warstwie ukrytej. Przez to sprawdza się w sytuacjach, gdzie nieznane jest całe środowisko, zachowując sprawność w momentach gdzie wymagane jest planowanie.

4. Przegląd literatury:

Przydatne repozytoria:

implementacje MuZero:

<https://github.com/johan-gras/MuZero>

<https://github.com/opensdilab/LightZero>

<https://github.com/werner-duvaud/muzero-general>

Zbiór gier atari:

<https://github.com/Farama-Foundation/Arcade-Learning-Environment>

<https://paperswithcode.com/dataset/dqn-replay-dataset>

Potencjalne inne środowiska do uczenia:

<https://github.com/clvrai/awesome-rl-envs>

<https://github.com/Farama-Foundation/ViZDoom>

5. Implementacja

Wybrane narzędzia i biblioteki: Python, torch

Wybrane środowisko gry: Atari 60 games

Baseline, metryki: wyniki eksperymentów w artykule

Zaproponowane modyfikacje: DOOM

Podział zadań: Marek – pisanie dokumentacji + trening/testy + integracja z grami

Paulina – implementacja modelu i pierwsze testy + wsparcie

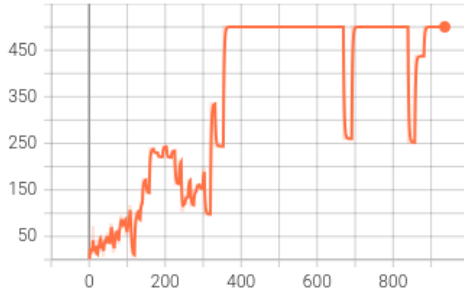
6. Plan eksperymentu

Spróbować nauczyć wybrane środowiska tak, aby osiągnęły zadowalające wyniki (przejście gry/przetrwanie/wysoki wynik).

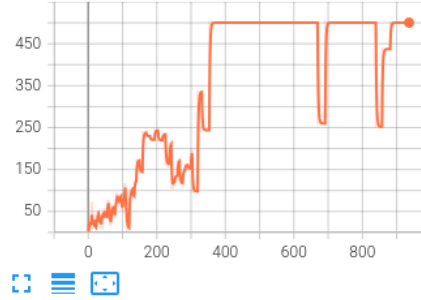
7. Wyniki

Cartpole

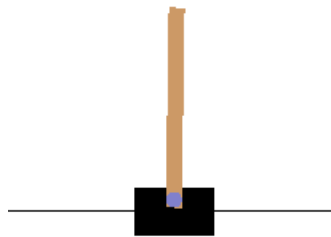
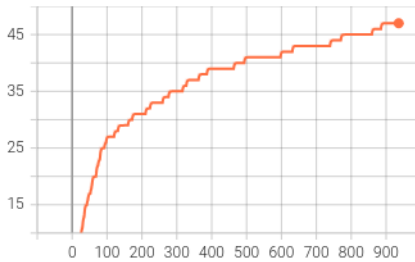
1.Total_reward/1.Total_reward
tag: 1.Total_reward/1.Total_reward



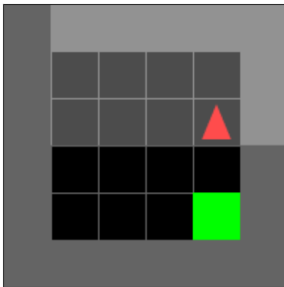
1.Total_reward/3.Episode_length
tag: 1.Total_reward/3.Episode_length



2.Workers/1.Self_played_games
tag: 2.Workers/1.Self_played_games

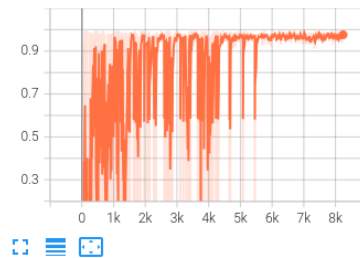


Gridworld

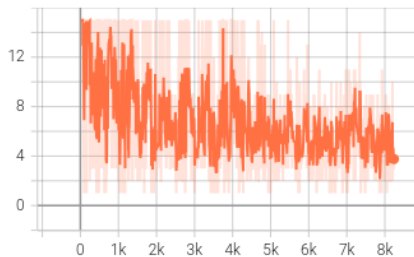


get to the green goal square

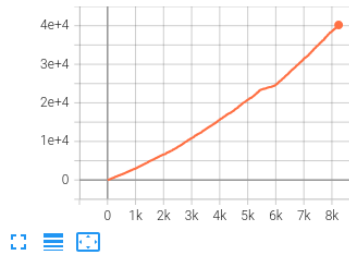
1.Total_reward/1.Total_reward
tag: 1.Total_reward/1.Total_reward



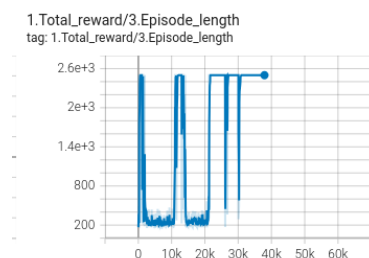
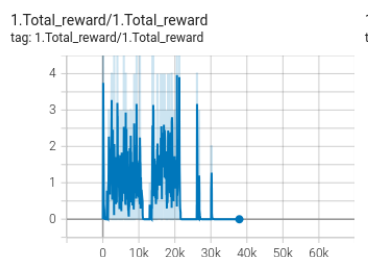
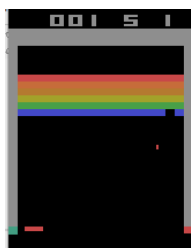
1.Total_reward/3.Episode_length
tag: 1.Total_reward/3.Episode_length



2.Workers/1.Self_played_games
tag: 2.Workers/1.Self_played_games



Breakout



Tutaj prezentacje :)

8. Bibliografia

<https://arxiv.org/pdf/1911.08265>

<https://github.com/johan-gras/MuZero>

<https://github.com/opensdilab/LightZero>

<https://github.com/werner-duvaud/muzero-general>

<https://github.com/Farama-Foundation/Arcade-Learning-Environment>

<https://paperswithcode.com/dataset/dqn-replay-dataset>

<https://github.com/clvrai/awesome-rl-envs>

<https://github.com/Farama-Foundation/ViZDoom>

<https://paperswithcode.com/sota/atari-games-100k-on-atari-100k>

