

# Trenowanie wirtualnego agenta w grze FPS z wykorzystaniem uczenia ze wzmocnieniem



Prezentacja tematu pracy dyplomowej: **Rafał Nowak**

# Wprowadzenie

Do tematu pracy zainspirowała mnie historia komputera Deep Blue ~ komputera grającego w szachy stworzonego przez firmę IBM i historia jak ówczesny mistrz świata Garii Kasparow 10-tego lutego 1996 roku wygrał przeciwko maszynie. Był to temat w tamtym okresie, który obiegił świat. Osoby w tamtym okresie, które nigdy nie grały w szachy, ekscytowały się tym niecodziennym pojedynkiem.

Porównywanie ludzi do maszyn nie jest niczym zupełnie nowym i tutaj narodził się mój pomysł na pracę inżynierską gdzie będę porównywał swoje osiągnięcia w grach typu FPS(strzelanki pierwszoosobowe ang. First Person Shooter) do osiągnięć stworzonego przeze mnie agenta za pomocą różnych, bliżej nieokreślonych algorytmów.

## Na czym będzie polegało porównanie?

To będzie pierwszy problem do rozwiązania w mojej pracy, czyli stworzenie aplikacji, która będzie w stanie porównać moje wyniki w zakresie precyzji i czasu reakcji a stworzonego przeze mnie agenta. Istnieją aplikacje dzięki którym można poprawić swoją celność w grach FPS i wyćwiczyć tzw. „pamięć mięśniową”. Jedną z takich aplikacji jest Kovaak 2: The Meta, którą będę starał się odwzorować na silnikach graficznych Unity i Unreal Engine – by jednocześnie porównać działanie algorytmu uczącego w różnych środowiskach.

# Na czym będzie polegało porównanie?

Porównanie ze sztuczną inteligencją będę przeprowadzał w grze nazwanej „Reflex Flick” polega ona na strzelaniu w punkty losowo pojawiające się na ekranie, a głównym celem jest zestrzelenie jak największej ilości punktów w określonym czasie (1 minuta). Punkty pojawiają się co sekundę, na ok. 400ms





Stworzenie sztucznej inteligencji za pomocą wzmocnienia wiąże się z kilkoma odwiecznymi problemami z którymi będę musiał się zmierzyć podczas implementacji. Mogłoby się wydawać, że nauka sztucznej inteligencji w ten sposób jest pozbawiona wad, nic bardziej mylnego.



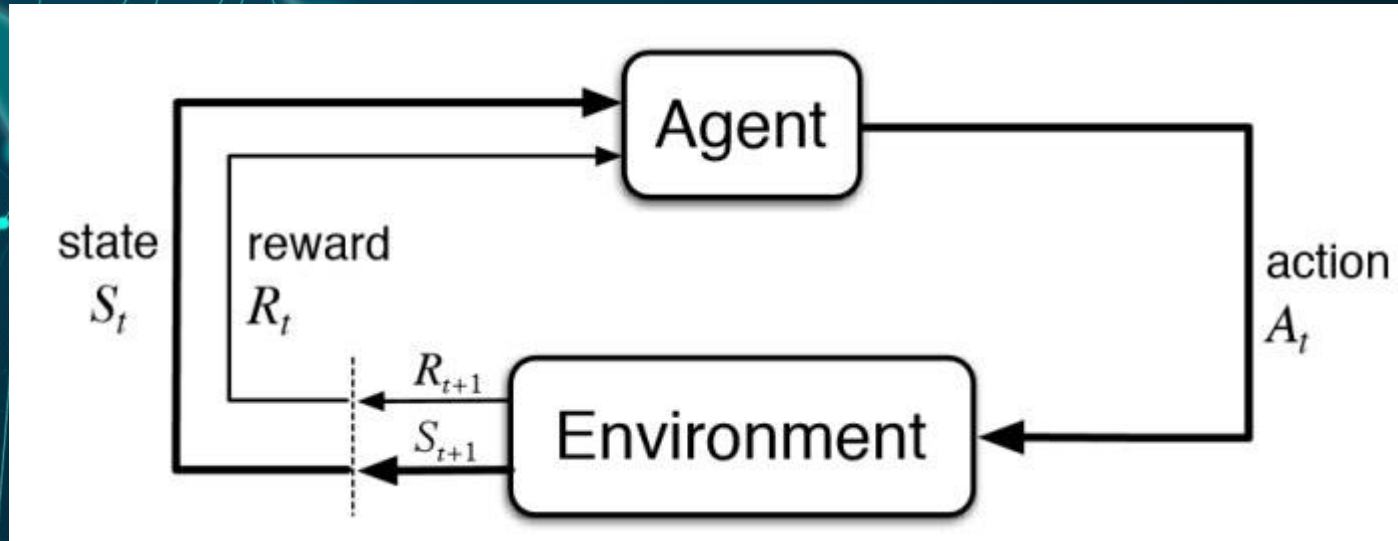
**AI**

# Uczenie ze wzmocnieniem

Uczenie ze wzmocnieniem nie jest tworzone na gotowych danych tak jak uczenie nadzorowane. W uczeniu nadzorowanym podajemy naszej sztucznej inteligencji dane dzięki, którym się uczy. Problem w uczeniu nadzorowanym jest taki że nigdy nie będzie lepsza niż dane, które dostała np. jeśli podamy sztucznej inteligencji dane najlepszego gracza szachowego, to nigdy nie będzie od niego lepsza. Zamiast tego nagradzamy naszą sztuczną inteligencję za uzyskanie jak najlepszego wyniku lub osiągnięcie określonego celu.



# Uczenie ze wzmocnieniem



# Problemy

01

## Credit assignment problem

Nagradzanie sztucznej inteligencji za wygraną a nie za kroki, które do tej wygranej doprowadziły

03

## The alignment problem

AI okazuje się na tyle sprytne, że tworzy rozwiązanie, które go nagradzają, ale niekoniecznie w taki sposób jakibyśmy chcieli.

02

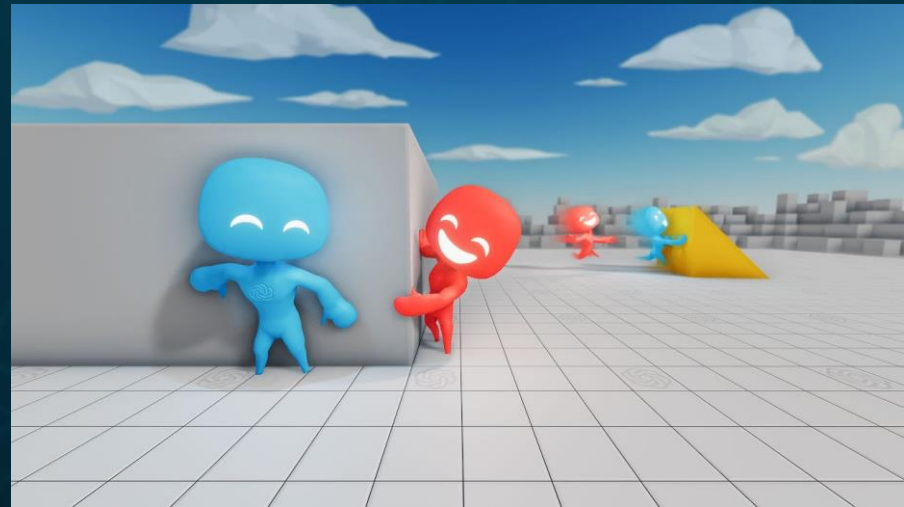
## Reward shaping

Jest to nagradzanie AI za poszczególne kroki jakie musi podjąć by osiągnąć zamierzany efekt, problem polega na tym, że ręcznie trzeba konfigurować AI za każdym razem gdy zmienia się środowisko



# Uczenie ze wzmocnieniem

Uczenie ze wzmocnieniem często prowadzi do nieprzewidywalnych rozwiązań, przykładem tego jest eksperymentalna gra w chowanego, gdzie jedni agenci wcielali się w rolę chowających a druga drużyna w szukających. AI została przyuczona miliony razy i cały czas jedna z drużyn znajdowała rozwiązanie by pokonać drugą drużynę – wykorzystując, również błędy w grze. Zachęcam do zapoznania się z materiałem:  
<https://youtu.be/Lu56xVlZ40M>



# Dziękuję za uwagę

Ewentualne pytania proszę pisać na:

[nr41460@zut.edu.pl](mailto:nr41460@zut.edu.pl)

Autor:

Rafał Nowak

CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, and infographics & images by **Freepik**.

Promotor projektu:

Dr. Inż. Joanna Kołodziejczyk