

# Tvorba umelej inteligencie pre počítačové hry



Riešitelia: Bc. Michal Jakabovič, Bc. Jakub Janky, Bc. Matúš Martinčo, Bc. Dávid Matis

Vedúci práce: Ing. Peter Tarábek, PhD., Ing. Marek Baláž

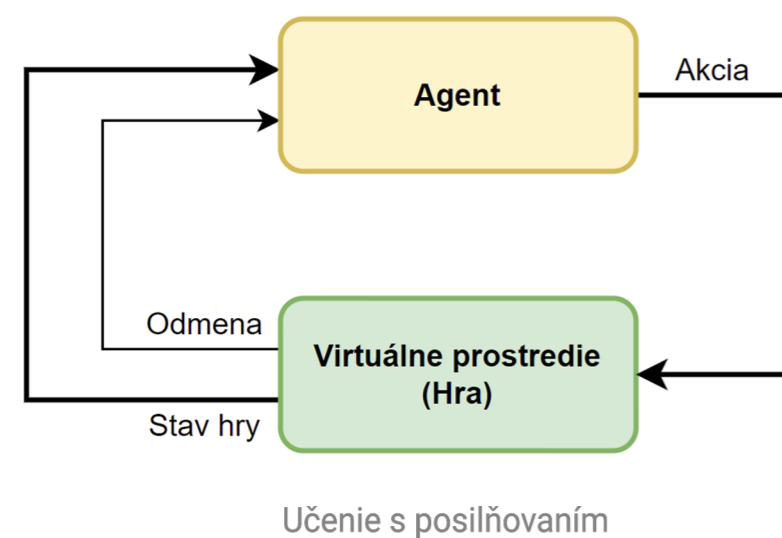
ŽILINSKÁ UNIVERZITA V ŽILINE  
Fakulta riadenia  
a informatiky

## Projekt



Projekt sa zaoberá tvorbou umelej inteligencie pre počítačové hry využitím **reinforcement learning-u** (učenie s posilňovaním).

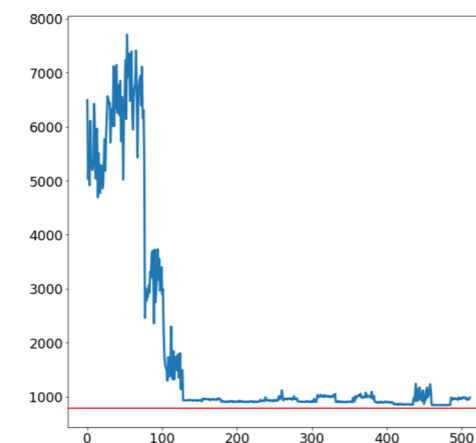
V praxi má Reinforcement Learning široké využitie a tento projekt poskytuje študentom, zaujímavou a zábavnou formou, možnosti naučiť sa, ako funguje a ako je ho možné efektívne aplikovať.



## Úloha p-medián

Zaoberali sme sa možnosťami riešenia **optimalizačných úloh** pomocou RL. Táto téma je zaujímavá, keďže mnohé hry sa podobajú kombinatorickým optimalizačným úlohám. Ako zástupcu tohto problému bola vybratá úloha **p-medián**. Do budúcnosti máme v pláne preskúmať ďalšie potenciálne prístupy k riešeniu týchto úloh.

- Bolo vytvorené prostredie s ktorým sú agenti schopní pracovať (jediná možná akcia je pridávať - "postaviť sklad")
- Malé experimenty s 5 [p=4] kandidátmi a 20 [p=10] kandidátmi
- Riešene pomocou Q Tabuľky a PPO
- Experimentovanie s parametrami počas tréningu



Vývoj účelovej funkcie počas tréningu PPO (červená - optimálne riešenie); 20 kandidátov, p=10, 100 000 hier

## Motivácia



Oboznámenie sa so základnými princípmi učenia s posilňovaním

Myšlienka Q-learning-u

Zapojenie (konvolučných) neurónových sietí

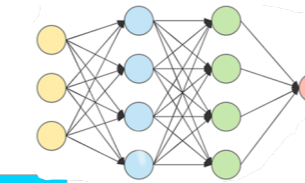
Implementácia, aplikácia a tréning algoritmov založených na reinforcement learning-u vo vybraných herných prostrediach

Skúmanie možných úprav algoritmov pre riešenie abstraktnejších problémov

Q - Learning

stavy	akcie			
	↑	↓	←	→
(1,1)	0	0	0	0
(1,2)	0	0	0	0
...	...	...	...	...
(m,n)	0	0	0	0

Deep Q - Learning (DQN)



A2C

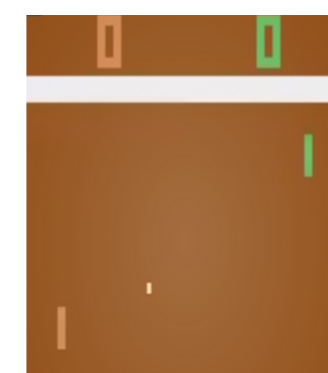
PPO

RND

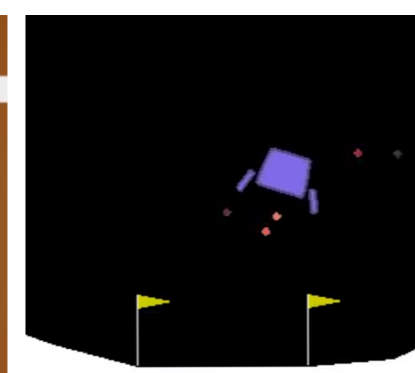
## OpenAI Gym



- Platforma, zameraná na výskum v oblasti učenia s posilňovaním, poskytujúca širokú škálu prostredí
- Jednoduché počiatkové prostredie s ktorým sme začínali pre nadobudnutie skúseností. Boli vykonané mnohé malé experimenty pre porovnanie prístupov k tréningu na prostrediach **Lunar Lander** a **Pong**



Pong - v0

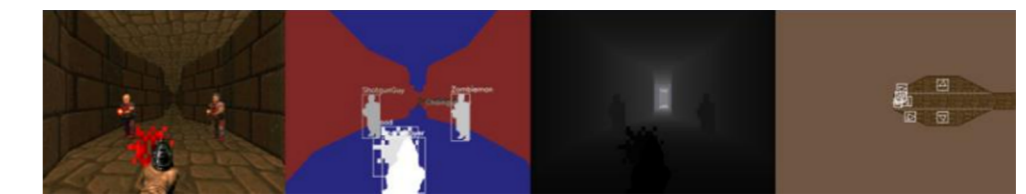


Lunar Lander - v2

## ViZDoom



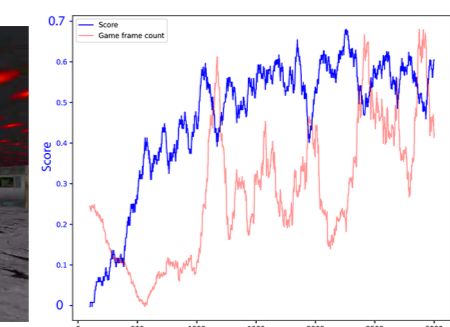
- Pokročilá platforma na vývoj botov v hre Doom s podporou synchronných a asynchronných single-player a multi-player módov
- K dispozícii sú informácie z hry v podobe obrazu



V práci sme sa zamerali hlavne na úlohy:

- Predikcia pozície nepriateľa (scenár Predict position)
- Obrana voči vlnám nepriateľov (scenár Defend the Center)

Predict position



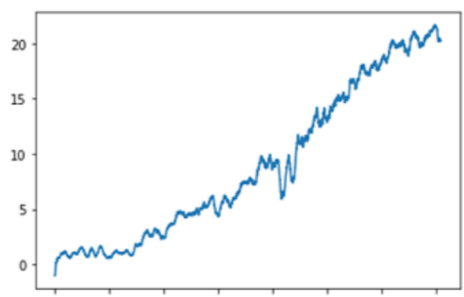
Predict position - priemerná úspešnosť (modrá) počas tréningu DQN agenta (teoretické maximum = 1)  
Pre zefektívnenie tréningu bola použitá suma budúcich odmien

Počas práce s Predict position boli použité viaceré spôsoby pre efektívne tréningovanie agenta:

- suma budúcich odmien - väčší výhľad do budúcnosti
- odmena priradená okamžite po výstrele
- kombinácia viacerých agentov - **Ensemble agent**

Defend the Center - priemerné skóre ~22 (teor. maximum = 25) počas tréningu A2C agenta. Dĺžka tréningu ~5 hodín

Bola využitá penalizácia za netrafenie nepriateľa



Defend the Center

## Grad-CAM



Výstup z metódy Grad-CAM zobrazený nad obrazom z hry v podobe teplotnej mapy

- Metóda slúžiaca na lepšiu vysvetliteľnosť konvolučných neurónových sietí
- Bola využitá pri scenári **Defend the Center**
- Vizualizácia oblastí, ktoré v najväčšej miere ovplyvnili výstup z neurónovej siete