

Posudek oponenta bakalářské práce

Student: Glós Kristián
Téma: Hluboké učení AI v herních prostředích (id 24101)
Oponent: Bobák Petr, Ing., UPGM FIT VUT

1. **Náročnost zadání** **obtížnější zadání**
Vzhledem k experimentálnímu charakteru je zadání práce **obtížnější**. Autor musel nastudovat složitou problematiku posilovaného učení a získané znalosti aplikovat při návrhu herního prostředí.
2. **Splnění požadavků zadání** **zadání téměř splněno**
Poslední bod zadání (prezentujte výsledky pomocí plakátu a krátkého videa) není splněn. Mírně diskutabilní je i druhá část bodu 3 (implementujte zvolené algoritmy a experimentujte s jejich parametry), jelikož algoritmy učení nebyly implementovány, nýbrž převzaty z volně dostupných knihoven.
3. **Rozsah technické zprávy** **je v obvyklém rozmezí**
Rozsahem se práce nachází **mírně pod hranicí obvyklého rozmezí** - závěr je uveden na straně 38 (cca 53 normostran).
4. **Prezentační úroveň předložené práce** **65 b. (D)**
Členění textu není ideální stejně jako pojmenování některých kapitol a sekcí, což v kombinaci s dalšími prezentačními nedostatky negativně ovlivňuje pochopení textu práce.

Níže je výčet nejpalcivějších nedostatků:

- V kapitole 4 autor prezentuje výsledky experimentů, bohužel však v celém textu práce **není uvedeno jakým algoritmem učení byly tyto výsledky získány** (např. DQN, PPO, SAC, TRPO?), jak vypadala použitá **architektura sítě** a další důležité **hyperparametry**. Odvozené poznatky jsou tak pro čtenáře téměř nicneříkající, jelikož chování agenta je zásadně ovlivněno všemi zmíněnými aspekty.
- V sekci 2.5 "Reinforcement Learning" autor přibližuje tabulkovou verzi metody Q-Learning a velmi stručně metodu PPO, nicméně **význam a role neuronových sítí v RL je zcela opominut**.
- Sekce 4.1 pojmenovaná "Analýza sítí" popisuje chování konkrétního algoritmu, o analýze rozličných architektur sítí není ani zmínka.
- V úvodu druhé kapitoly autor tvrdí, že naučení šachu tradičními algoritmy AI by trvalo více než historie celého vesmíru - toto tvrzení je však založeno na literatuře z roku 1950.
- Obrázky v kapitole 3 a 4 nejsou dostatečně popsány a často z nich není patrné co přesně obrázek ilustruje.
- Do sekce 2.4 "Hlboké učenie" jsou nelogicky zanořeny i přístupy učení - s/bez učitele, aj. Obdobně neobratně je členěna i sekce 3.7 "Development".
- Kapitola 2 je nepříliš vhodně pojmenovaná jako "Teória".

V textu se **zřídka vyskytují také faktické chyby** (např. nepřesná definice eps-greedy, str. 14, má být argmax) případně vágní/nejednoznačné vyjadřování:

- "... *algoritmy* ako sú Reinforcement Learning a Neuroevolution, ci taktiež Imitation Learning..." (RL/IL nelze označit za algoritmy)
- "...*trojrocná* neurónová sieť Leela Chess Zero dokázala porazit [8] *12 ročný* tradičný chess..."
- discount factor je jednou označen jako gamma, v pozdější části textu jako γ

5. **Formální úprava technické zprávy** **75 b. (C)**
Práce je psána ve slovenském jazyce. Po pravopisné stránce se v textu vyskytuje **minimum chyb a překlepů**. Autor v textu hojně využívá **autorský plurál**, což především v popisu implementace působí nepatřičně (např. "V *našej bakalárke* sme na to narazili mnohokrát ...", "... algoritmy sme po oboznámení sa implementovali ...")

Po typografické stránce je formátování textu **na průměrné úrovni** - rovnice nejsou číslovány, některé obrázky nejsou odkazovány z textu, grafy vizualizující metriky z trénování jsou jen printscreenem z TensorBoardu, chybí popisy os, popis obrázku 3.8 duplikuje popis obrázku 3.7, aj.

6. **Práce s literaturou** **90 b. (A)**

Autor cituje **28 kvalitních a relevantních zdrojů** souvisejících s řešeným problémem včetně **osmi vědeckých studií a jedné knihy**. Všechny elektronické zdroje jsou citovány ve stejný den těsně před termínem odevzdání BP.

7. Realizační výstup**75 b. (C)**

Výstupem práce jsou dvě herní prostředí vhodné pro experimenty s posilovaným učením. V prvním agent musí postupně splnit několik menších úkolů jejichž splněním se zpřístupní hlavní cíl prostředí. Druhé prostředí je bludiště, ve kterém je úkolem nalézt cestu k cíli. Obě prostředí jsou implementovány v Unity, pro učení agenta byly použity metody z knihovny ML-Agents a UnityNEAT. Následně autor zkoumal vliv různých parametrů těchto prostředí a experimentoval s vícero přístupy učení. Výsledkem experimentů s parametry prostředí jsou poznatky o vlivu složitosti, míře náhodnosti, počtu podúkolů, aj. Získané poznatky jsou však z velké části specifické pouze pro dvě implementovaná prostředí a nelze je zobecnit i na další úlohy. Přesvědčivosti bohužel nenahrává ani zcela chybějící popis parametrů jednotlivých experimentů - není jasné jakými algoritmy bylo dosaženo prezentovaných výsledků, jaká byla architektura neuronové sítě a další hyperparametry.

8. Využitelnost výsledků

Vytvořené dílo nenavazuje na žádnou předchozí práci či výzkum. Implementačně práce nepřináší mnoho nového, je však dobrým základem pro další výzkum.

9. Otázky k obhajobě

1. Objasněte prosím počáteční podmínky experimentů uvedených v sekci 4.1. Jakými algoritmy bylo dosaženo prezentovaných výsledků? Jak vypadala architektura neuronové sítě? Co reprezentují vstupy a výstupy sítě? Experimentoval jste s hyperparametry?
2. Vysvětlíte prosím co ilustrují obrázky 4.1(a) a (b). Co reprezentují jednotlivé průběhy vizualizované oranžovou a modrou barvou?

10. Souhrnné hodnocení**78 b. dobře (C)**

Autor v rámci bakalářské práce prostudoval **složitou** problematiku **posilovaného učení, imitačního učení a neuroevoluce**. Navržené prostředí i provedené experimenty jsou dobrým odrazovým můstkem pro další výzkum. Vzhledem k chybějícímu popisu experimentů jsou však jejich výsledky méně přesvědčivé.

Prohlášení: Uděluji VUT v Brně souhlas ke zveřejnění tohoto posudku v listinné i elektronické formě.

V Brně dne: 28. května 2021

Bobák Petr, Ing.
oponent