

**I. IDENTIFIKAČNÍ ÚDAJE**

<b>Název práce:</b>	<b>Porovnání standardních metod posilovaného učení</b>
<b>Jméno autora:</b>	<b>Michaela Cihlářová</b>
<b>Typ práce:</b>	bakalářská
<b>Fakulta/ústav:</b>	Fakulta elektrotechnická (FEL)
<b>Katedra/ústav:</b>	Katedra Kybernetiky FEL ČVUT
<b>Vedoucí práce:</b>	Karel Zimmermann
<b>Pracoviště vedoucího práce:</b>	Katedra Kybernetiky FEL ČVUT

**II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ**

<b>Zadání</b>	<b>náročnější</b>
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
Hlavní motivací této práce byl vznik nové studentské úlohy pro předmět o učení hlubokých neuronových sítí. Vyřešení vyžaduje kombinaci různých dovedností jako je např. (i) vymodelování simulačního prostředí v Blenderu, (ii) ozkoušení existujících implementací metod posilovaného učení, (iii) naimplementování vybrané metody, (iv) vytvoření rozumného zadání, obsahujícího template kódu a formální popis implementační práce a experimentů, které mají studenti provést.	

<b>Splnění zadání</b>	<b>splněno s menšími výhradami</b>
<i>Posuďte, zda předložená závěrečná práce splňuje zadání. V komentáři případně uveďte body zadání, které nebyly zcela splněny, nebo zda je práce oproti zadání rozšířena. Nebylo-li zadání zcela splněno, pokuste se posoudit závažnost, dopady a případně i příčiny jednotlivých nedostatků.</i>	
<ul style="list-style-type: none"> <li>(i) Povedlo se vytvořit prostředí kompatibilní s Algym, ikdyž stav neobsahuje kamerové-lidarové obrázky (po dohodě se studentkou nepovažuji za závažný problém).</li> <li>(ii) Bylo ozkoušeno několik metod, ale není jasná architektura sítí. Mohla být provedena rozsáhlejší studie vlivu parametrů na jednotlivé metody.</li> <li>(iii) Re-implementace A2C je správná, dosahuje výsledků srovnatelných se stable baselines.</li> <li>(iv) Žádné zadání nabylo vytvořeno.</li> </ul>	

<b>Aktivita a samostatnost při zpracování práce</b>	<b>B - velmi dobře</b>
<i>Posuďte, zda byl student během řešení aktivní, zda dodržoval dohodnuté termíny, jestli své řešení průběžně konzultoval a zda byl na konzultace dostatečně připraven. Posuďte schopnost studenta samostatné tvůrčí práce.</i>	
Studentka se pravidelně účastnila schůzek, ale postup prací byl občas pomalejší.	

<b>Odborná úroveň</b>	<b>C - dobře</b>
<i>Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	
Není jasný vliv zvolené architektury policy a hyper-parametrů u testovaných metod. Experimenty zmiňují „trojvrstvou lineární síť“, co to je a kolik má parametrů, je to jediná testovaná architektura?	

<b>Formální a jazyková úroveň, rozsah práce</b>	<b>C - dobře</b>
<i>Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.</i>	
Práce byla napsána „na poslední chvíli“ a nezbyl čas na iterování textu => V zápisech je mnoho nepřesností (viz níže).	
<ul style="list-style-type: none"> <li>1. Rovnice (2.2): Je neobvyklé, že reward není funkcí akce. Matematicky není korektní sum přes <math>r_t</math> z <math>\tau</math>, asi spíš mělo být <math>(s_t, a_t)</math> z <math>\tau</math>.</li> <li>2. Rovnice (2.10): Použití <math>\tau</math> je nepřesné, není jasné v jakém bodě trajektorie začíná, navíc <math>r(\tau)</math> je nedefinováno.</li> <li>3. Citace A2C na str 8 je špatně</li> <li>4. Rovnice (2.26): <math>\tau</math> nedefinováno</li> </ul>	

5. Rovnice (2.26):  $\theta_{pi}$  a funkce P nedefinováno. Navíc není jasné co znamená „stačí zadat  $L_{(\theta_V)=...}$ “. M2 to jasno je, ale jakémukoli jinému čtenáři jen těžko.
6. Obr 4.4: vzdálenost má být dist, odměna má být dist\_rew
7. Obr 4.5: podobně neodpovídá rovnicím
8. Obr 5.3: není jasné, čemu odpovídají jednotlivé barvy a proč jsou u každé metody jiné. Stejný scale obrázků by výrazně pomohl v porovnání jednotlivých metod.
9. Občas chyby slova: „V této verzi se Actor a Critic až po pěti krocích simulace.“

**Výběr zdrojů, korektnost citací**

**A - výborně**

*Vyjádřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení závěrečné práce. Charakterizujte výběr pramenů. Posuďte, zda student využil všechny relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků a úvah, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami.*

Studentka citovala práce navržené v zadání.

**Další komentáře a hodnocení**

*Vyjádřete se k úrovni dosažených hlavních výsledků závěrečné práce, např. k úrovni teoretických výsledků, nebo k úrovni a funkčnosti technického nebo programového vytvořeného řešení, publikačním výstupům, experimentální zručnosti apod.*

Vložte komentář (nepovinné hodnocení).

**III. CELKOVÉ HODNOCENÍ A NÁVRH KLASIFIKACE**

*Shrňte aspekty závěrečné práce, které nejvíce ovlivnily Vaše celkové hodnocení.*

Studentka se aktivně účastnila schůzek, vytvořila simulační prostředí kompatibilní Algym a re-implementovala vybranou metodu A2C, jejíž chování je srovnatelné se stable baselines. Text práce je spíše slabší, obsahuje mnoho drobných chyb a neúplných vysvětlení. Hlavní motivací této práce byl vznik nové úlohy do předmětu o učení hlubokých sítích, žádné zadání však nevzniklo (přestože jsme o něm při pravidelných schůzkách hovořili).

Otázky do diskuze:

- (i) Má reward funkce lokální extrémy? Jak by vypadal, třeba řez YZ-rovinou?
- (ii) Jaké architektury policy se použili u A2C, CMA-ES a jiných testovaných metod?

Předloženou závěrečnou práci hodnotím klasifikačním stupněm **C - dobře**.

Datum: 3.6.2021

Podpis:

**I. IDENTIFIKAČNÍ ÚDAJE**

<b>Název práce:</b>	<b>Porovnání standardních metod posilovaného učení</b>
<b>Jméno autora:</b>	<b>Michaela Cihlářová</b>
<b>Typ práce:</b>	bakalářská
<b>Fakulta/ústav:</b>	Fakulta elektrotechnická (FEL)
<b>Katedra/ústav:</b>	Katedra kybernetiky
<b>Oponent práce:</b>	Ing. Petr Pošík, Ph.D.
<b>Pracoviště opONENTA práce:</b>	Katedra kybernetiky, FEL ČVUT

**II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ**

<b>Zadání</b> <i>Hodnocení náročnosti zadání závěrečné práce.</i>	<b>náročnější</b>
Porozumět variantám RL algoritmů a jednotlivým rozdílům v jejich principech považuji za náročnější část.	

<b>Splnění zadání</b> <i>Posuďte, zda předložená závěrečná práce splňuje zadání. V komentáři případně uveďte body zadání, které nebyly zcela splněny, nebo zda je práce oproti zadání rozšířena. Nebylo-li zadání zcela splněno, pokuste se posoudit závažnost, dopady a případně i příčiny jednotlivých nedostatků.</i>	<b>splněno</b>
Všechny body zadání považuji za splněné, ačkoli nejsem schopen plně posoudit, zda jsou výsledky práce využitelné jako domácí úloha pro studenty předmětu VIR.	

<b>Zvolený postup řešení</b> <i>Posuďte, zda student zvolil správný postup nebo metody řešení.</i>	<b>správný</b>
Postup se velmi přesně shoduje s pokyny pro vypracování.	

<b>Odborná úroveň</b> <i>Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	<b>C - dobře</b>
Studentka v práci využila znalosti především z předmětu KUI, které musela doplnit o poznatky z odborné literatury. Studentka se dobře vypořádala s náročností teoretické části. V textu práce na mnoha místech užívá slovo „sít“, ačkoli daná funkce nemusí být obecně reprezentovaná neuronovou sítí. V práci mi ovšem chybí bližší popis knihovny stable-baselines. Není mi např. jasné, zda jednotlivé algoritmy z této knihovny mají předdefinované metaparametry modelů (např. rozměry sítě), které jednotlivé algoritmy trénují, nebo zda je uživatel volí sám. Pokud je volí sám, postrádám popis, jak byly zvoleny. Podobně u algoritmu CMA-ES mi není zcela jasné, co vlastně algoritmus ladil a jak byl do celého RL procesu zapojen. Některé použité termíny nejsou přesně definované (např. trajektorie).	

<b>Formální a jazyková úroveň, rozsah práce</b> <i>Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.</i>	<b>B - velmi dobře</b>
Práce je po typografické stránce zdařilá, členění práce je logické. Po jazykové stránce není práce perfektní, na několika místech text nedává dobrý smysl, nebo např. chybí slovo. Podobně jsou v práci jisté nedostatky v použití matematické notace, např. v rovnici (2.7) funkce argmax vrací jednu akci, nikoli celou policy, v rovnicích (2.29) a (2.30) je výsledkem funkce argmax sada parametrů $\theta$ , nikoli nějaké L, jehož význam jsem také nenašel.	

<b>Výběr zdrojů, korektnost citací</b> <i>Vyjádrete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení závěrečné práce. Charakterizujte výběr pramenů. Posuďte, zda student využil všechny relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků a úvah, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami.</i>	<b>A - výborně</b>
---	--------------------

Citované zdroje jsou relevantní. Nezaznamenal jsem porušení citační etiky. V sekci 2.5.1 je patrně chybný odkaz na zdroj k algoritmu A2C.

#### **Další komentáře a hodnocení**

*Vyjádřete se k úrovni dosažených hlavních výsledků závěrečné práce, např. k úrovni teoretických výsledků, nebo k úrovni a funkčnosti technického nebo programového vytvořeného řešení, publikačním výstupům, experimentální zručnosti apod.*

### **III. CELKOVÉ HODNOCENÍ, OTÁZKY K OBHAJOBĚ, NÁVRH KLASIFIKACE**

*Shrňte aspekty závěrečné práce, které nejvíce ovlivnily Vaše celkové hodnocení. Uvedte případné otázky, které by měl student zodpovědět při obhajobě závěrečné práce před komisí.*

Na práci velmi kladně hodnotím, že zde došlo k propojení několika technologií (Blenderu pro tvorbu modelů, PyBulletu pro fyzikální simulaci, aigymu pro benchmarkování RL algoritmů) a znalostí získaných během studia. Nicméně v práci postrádám popis několika poměrně důležitých částí, díky čemuž je práce méně srozumitelná, než by mohla být.

Předloženou závěrečnou práci hodnotím klasifikačním stupněm **C - dobře**.

Mám na studentku následující doplňující otázky:

1. V sekci 4.3.1 definujete 20D stavový prostor a 12D akční prostor. Předpokládám, že výsledkem učení má být funkce mapující stavy na akce. Jak je tato funkce v práci reprezentována? Pokud neuronovou sítí, jaké jsou její metaparametry?
2. Vyjasněte, jak byl algoritmus CMA-ES zapojen do procesu RL. Co přesně ladil?
3. V závěru uvádíte, že jste pro implementaci v PyTorch zvolila algoritmus A2C, ačkoli TRPO dává lepší výsledky, kvůli nižší výpočetní náročnosti, aby úloha šla řešit na počítačích studentů. Tento aspekt ve své práci ale nijak neměříte, nediskutujete. Můžete rozdíl ve výpočetní náročnosti nějak kvantifikovat, např. měřením, o kolik je učení TRPO pomalejší než A2C?

Datum: 27.5.2021

Podpis: