

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

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

<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>	
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 „sítě“, 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>	<b>B - velmi dobře</b>
<i>Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.</i>	
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 $\operatorname{argmax}$ vrací jednu akci, nikoli celou policy, v rovnicích (2.29) a (2.30) je výsledkem funkce $\operatorname{argmax}$ sada parametrů $\theta$ , nikoli nějaké $L$ , jehož význam jsem také nenašel.	

<b>Výběr zdrojů, korektnost citací</b>	<b>A - výborně</b>
<i>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.</i>	

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. Uveďte 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: