

**I. IDENTIFIKAČNÍ ÚDAJE**

<b>Název práce:</b>	Reinforcement Learning for Swarm Control of Unmanned Aerial Vehicles
<b>Jméno autora:</b>	Karel Poncar
<b>Typ práce:</b>	bakalářská
<b>Fakulta/ústav:</b>	Fakulta elektrotechnická (FEL)
<b>Katedra/ústav:</b>	Katedra kybernetiky
<b>Oponent práce:</b>	doc. Ing. Zdeněk Hurák, Ph.D.
<b>Pracoviště oponenta práce:</b>	Katedra řídicí techniky FEL ČVUT

**II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ**

<b>Zadání</b>	<b>průměrně náročné</b>
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
S vědomím, že projekty využívající různé (cizí) softwarové knihovny často vyžadují velký objem práce, se domnívám, že šlo o zadání s adekvátní (průměrnou) náročností.	

<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 byly, alespoň jak můžu soudit, splněny.	

<b>Zvolený postup řešení</b>	<b>správný</b>
<i>Posuďte, zda student zvolil správný postup nebo metody řešení.</i>	
Nemám výhrad k postupu práce.	

<b>Odborná úroveň</b>	<b>B - velmi dobře</b>
<i>Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	
Domnívám se, že úroveň odbornosti je velmi dobrá. Níže (zejména v závěru) uvádím konkrétní komentáře.	

<b>Formální a jazyková úroveň, rozsah práce</b>	<b>A - výborně</b>
<i>Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.</i>	
Práce je psána srozumitelnou angličtinou. Vysázena je pomocí LaTeXu, a tak je výsledný dojem velmi solidní. Nicméně několik formálních zlepšení můžu (pro příště) doporučit:	
<ul style="list-style-type: none"> <li>• Popisky os některých obrázků (např. Obr. 4.3 či 4.6) jsou tak malé, že jsou jen velmi těžce čitelné. Coby čtenář můžu nabýt dojmu, že autorovi vůbec nezáleželo na tom, zda informaci předanou obrázkem spolehlivě přijmu, a že možná je tam obrázek jen coby dekorace...</li> <li>• V angličtině se při použití desetinné tečky (místo čárky) pak coby symbol pro násobení už spíše radši nepoužívá tečka, aby nedošlo ke zmatení. Tedy místo <math>1.2 \cdot 10^3</math> radši <math>1.2 \times 10^3</math>. Viz například Thompson, A., and B. N. Taylor. 'Guide for the Use of the International System of Units (SI)', 30 March 2008. <a href="https://www.nist.gov/publications/guide-use-international-system-units-si">https://www.nist.gov/publications/guide-use-international-system-units-si</a>.</li> <li>• Fyzikální jednotky (v práci například m pro metry) mají být sázeny zásadně vpřimeným fontem, viz například Beccari, Claudio. 'Typesetting Mathematics for Science and Technology According to ISO 31/XI'. <i>TUGboat</i>, Tutorials Surveys, 18, no. 1 (1997): 39-47. <a href="https://tug.org/TUGboat/tb18-1/tb54becc.pdf">https://tug.org/TUGboat/tb18-1/tb54becc.pdf</a>. Mimo jiné se tak předejde konfliktu se proměnnými (i v práci se objevuje proměnná <math>m</math>). V LaTeXu možno použít balíček <i>siunitx</i> <a href="https://ctan.org/pkg/siunitx">https://ctan.org/pkg/siunitx</a>.</li> </ul>	

<b>Výběr zdrojů, korektnost citací</b>	<b>C - dobře</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</i>	

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.

V práci autor tvrdí, že v oblasti řízení rojů bezpilotních prostředků pomocí RL nebylo dosud prováděno dost výzkumu. V přehledu literatury se pak odkazuje na jeden článek o řízení jedné kvadrokoptéry a jeden článek (napsaný jeho vedoucím práce) o řízení roje dronů. Je tomu ale skutečně tak, že žádný jiný relevantní výzkum v této oblasti neexistuje? Po zadání klíčových slov jako „reinforcement learning swarm quadrotors“ lze v Google Scholar dostat nekonečné seznamy článků s relevantně znějícími názvy. Určitě nebylo nutné u bakalářské práce vytvářet naprosto vyčerpávající přehled stávajícího stavu (to samo o sobě je práce na několik měsíců), nicméně takto to působí, že ten přehled nebyl získán vůbec žádný.

#### **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.

Při zavádění základních konceptů a pojmů z oblasti RL v sekci 3.1.1 doporučuji věnovat pozornost rozdílu mezi pojmy *reward* a *return*. Obávám se, že v práci použitá terminologie, kdy *value function* přiřazuje danému stavu *expected reward*, je nestandardní. Alespoň ve srovnání s tím, jak jsou tyto zavedeny v [bibli RL od Suttona a Barto](#). Tam v sekci 3.3 píšou: "If the sequence of rewards received after time step  $t$  is denoted  $R_{t+1}, R_{t+2}, R_{t+3}, \dots$ , then what precise aspect of this sequence do we wish to maximize? In general, we seek to maximize the expected return, where the return, denoted  $G_t$ , is defined as some specific function of the reward sequence. In the simplest case the return is the sum of the rewards:  $G_t = R_{t+1} + R_{t+2} + R_{t+3} + \dots + R_T$ , (3.7) where  $T$  is a final time step." Podobně i dále *discounted reward* vs *discounted return*. Domnívám se, že přichýlení se k zavedené terminologii je vhodné, protože zrovna v tomto případě je k nedorozumění blízko – maximalizace očekávaného "rewardu" (jak ho uvažují Sutton a Barto) vede na *greedy policy*...

Opět s odkazem na knihu od Suttona a Barto lze dále namítat, že pojmy v práci zmíněné coby údajně *nepřilíží běžné v MDP*, jako jsou posloupnosti trojic *state–action–reward* jsou ve skutečnosti naprosto běžné. Viz jejich sekce 3.1 (The Agent Environment Interface). Tam tomu říkají *trajektorie*.

### **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.

Co můžu posoudit na základě odevzdaného textu, studentem byla odvedena solidní práce. Přesto mám pár otázek a komentářů (u obhajoby nemusí být čteny všechny):

- Vlastně ani po přečtení práce nerozumím té celkové strukturu řídicího systému. Na jednu stranu je v práci uváděn fyzikální model, do kterého vstupem jsou požadované rychlosti otáčení jednotlivých motorů. Na druhou stranu pak (v sekci 4.2.2 a následně 4.2.3) je zmíněno, že je použit "body rate PID", který coby vstup bere vektor požadované úhlové rychlosti kvadrokoptéry. Ten PID byl autorem práce navrhován či aspoň laděn? Jestli ne, tak k čemu pak byl vlastně ten fyzikální model uváděn? Nestačilo pak už mít jen model toho výsledného zpětnovazebního propojení kvadrokoptéry a regulátoru úhlové rychlosti? Každopádně by v práci (či aspoň u obhajoby) bylo užitečné mít **přehledné blokové schéma, které by tu strukturu v práci navrhovaného řídicího systému ukazovalo**. Tedy co z měřených veličin je zpracováváno (v práci navrhovaným) řídicím systémem, a co je pak jeho výstupem – tedy akčním zásahem.
- Postrádám v práci stručné, přehledné a jednoznačné vymezení té mise, kterou mají společně kvadrotory splnit. Speciálně požadavek na *kohezi* je vysvětlován až v hloubce práce v sekci 4.2.4 o návrhu odměn, ale to už má čtenář za sebou dvě desítky stránek a pořád vlastně není zřejmé, jak se odlišuje úspěšné splnění mise od velmi úspěšného. Domnívám se, že u inženýrského projektu by to **vymezení požadavků (requirements/specifications) mělo být velmi explicitní a pohromadě na jednom místě někde na začátku práce**, vždyť je vlastně i nezávislé na volbě metody řešení.
- Jedním z hlavních náplní projektu byl vývoj simulátoru rojů kvadrokoptér. Vzhledem k populárnosti kvadrokoptér mezi amatéry i profesionály, v průmyslu i v akademickém světě se nabízí otázka, zda takový

simulátor už není k dispozici. Alespoň pro simulaci jedné kvadkoptéry, neboť taková je samozřejmě základní komponentou pro simulaci roje. Hledání odpovědi na takovou otázku nepochybně muselo být jedním z prvních kroků v takovémto projektu. Minimálně s motivací později výstupy vlastního simulátoru **srovnat s nějakým jiným simulátorem.**

- Podobným dotazem je, zda přece jen nebylo možné výsledky simulací **srovnat s experimentálními výsledky**, a to alespoň pro jeden dron. Vždyť jakou teď můžeme mít důvěru v reálnost výsledků simulací?
- Z možných algoritmů pro RL byl autorem vybrán algoritmus zvaný *Proximal policy optimization* (PPO). Není však ani naznačen důvod takové volby. Domnívám se, že důležitou kvalitou inženýrské (projektové) zprávy je **zdůvodnění** voleb a rozhodnutí.

Předloženou závěrečnou práci hodnotím klasifikačním stupněm **B - velmi dobře.** V závislosti na reakci studenta na komentáře připouštím, že adekvátní výsledná známka může být i o stupeň lepší či horší.

Datum: 7.6.2023

Podpis: