

**RADA NAUKOWA DYSCYPLINY
INFORMATYKA TECHNICZNA I TELEKOMUNIKACJA POLITECHNIKI WARSZAWSKIEJ**

zaprasza na
OBRONĘ ROZPRAWY DOKTORSKIEJ

mgr. inż. Konrada Godlewskiego

która odbędzie się we **wtorek, 9 maja 2023 roku**, o godzinie **10:30** w trybie stacjonarnym.

Temat rozprawy:

„Monte Carlo Tree Search and Reinforcement Learning methods for multi-stage strategic card game”

Promotor: dr hab. inż. Bartosz Sawicki, prof. uczelni – Politechnika Warszawska

Recenzenci: dr hab. inż. Aleksandra Kawala-Sterniuk, prof. uczelni – Politechnika Opolska

dr hab. inż. Aleksandra Świetlicka – Politechnika Poznańska

dr hab. inż. Jarosław Kurek, prof. uczelni - SGGW w Warszawie

Obrona odbędzie się w sali nr 4 w gmachu Starej Kotłowni (teren główny Politechniki Warszawskiej).

Z rozprawą doktorską i recenzjami można zapoznać się w Czytelni Biblioteki Głównej Politechniki Warszawskiej, Warszawa, Plac Politechniki 1.

Streszczenie rozprawy doktorskiej i recenzje są zamieszczone na stronie internetowej: <https://bip.pw.edu.pl/Postepowania-w-sprawie-nadania-stopnia-naukowego/Doktoraty/Wszczete-po-30-kwietnia-2019-r/Dyscyplina-informatyka-techniczna-i-telekomunikacja-dziedzina-nauk-inzynieryjno-technicznych/mgr-inz.-Konrad-Godlewski>

Przewodniczący Rady Naukowej Dyscypliny
Informatyka Techniczna i Telekomunikacja
Politechniki Warszawskiej
dr hab. inż. Jarosław Arabas, prof. uczelni

Streszczenie

Rozprawa doktorska

Metody Monte Carlo Tree Search oraz Reinforcement Learning w wieloetapowej strategicznej grze karcianej

mgr inż. Konrad Godlewski

Rozprawa przedstawia metody uczenia maszynowego wykorzystane do stworzenia agentów do gry karcianej "The Lord of the Rings: The Card Game" (LOTRCG). Cechą wyróżniającą grę jest nierywalizacyjny i złożony charakter rozgrywki. Gracz podejmuje decyzje w pięciu, ściśle określonych momentach podczas rundy. Akcje mają charakter sekwencyjny, czyli decyzja podjęta w jednej fazie gry wywiera wpływ na decyzję w etapie kolejnym. Taki mechanizm rundy wymaga strategicznego zarządzania kartami. Dodatkowo pomiędzy decyzjami występują zdarzenia losowe, które uniemożliwiają łączenie kilku momentów decyzyjnych.

LOTRCG jest grą typu collectible card game, co oznacza szeroki wybór kart dostępnych dla gracza. Karty posiadają różne statystyki oraz zdolności specjalne profilujące je do konkretnej fazy gry. Duża różnorodność kart przekłada się na dużą popularność tego tytułu wśród miłośników gier karcianych. Ten element decydujący o atrakcyjności gry stanowi zarazem duże wyzwanie dla agentów AI.

Złożony charakter rundy jest pierwszym obiektem badań tej rozprawy. W celu analizy poszczególnych faz gry zaimplementowano agenta strategii losowej. Uzyskane przy jego pomocy wyniki pozwoliły określić kluczowe momenty decyzyjne. Ponadto agent losowy posłużył do wykonywania losowego próbkowania drzewa algorytmu MCTS.

Rozprawa wykorzystuje dwie rodziny technik inteligencji obliczeniowej. Pierwszą z nich jest Monte Carlo Tree Search (MCTS). Metoda ta opiera się na wyszukiwaniu heurystycznym korzystającym z losowego próbkowania stanów gry. MCTS przechowuje rozgrywkę w postaci drzewa, które rozbudowywane jest w sposób iteracyjny. Podstawowa wersja algorytmu MCTS została rozbudowana o redukcję akcji, która pozwala na ograniczenie rozmiarów drzewa. Modyfikację tę oparto o wiedzę ekspercką, która pozwala na eliminację kart o niskiej użyteczności. Dodatkowo przeprowadzono optymalizację hiperparametrów MCTS nie tylko względem skuteczności, ale także pod kątem czasu obliczeniowego. W ten sposób uzyskano optymalny zestaw agentów MCTS, który osiągnął współczynnik 82.8% wygranych.

Drugą rodziną analizowanych technik jest uczenie ze wzmocnieniem (Reinforcement Learning - RL). Uczenie ze wzmocnieniem opiera się na metodzie prób i błędów, w którym agent jest w interakcji ze środowiskiem. Agent otrzymuje obserwację stanu gry, następnie wybiera akcję, która zostaje wykonana w środowisku. Środowisko odpowiada nagrodą oraz następną obserwacją. Celem agenta jest maksymalizacja sumy nagród na przestrzeni całego epizodu. Z uwagi na działanie agenta w środowisku o zmiennej liczbie akcji, selekcję algorytmów RL oparto na sposobie kodowania akcji. W pracy stworzono dwa typy kodowania. Pierwszy z nich to makroakcje, które pozwalają na uzyskanie stałej liczby akcji. Drugim typem są akcje bezpośrednie, czyli zagrywanie poszczególnych kart przez agenta. Dla makroakcji zaimplementowano algorytmy Q-Learning oraz Actor-Critic (AC). Q-Learning to szeroko stosowane rozwiązanie w dziedzinie RL aproksymujące wartość użyteczności stanu gry. Actor-Critic korzysta z aproksymacji zarówno funkcji użyteczności jak i strategii. Agent AC zaprogramowano do makroakcji jak i akcji bezpośrednich. W wyniku przeprowadzonej optymalizacji najlepszy zestaw agentów RL wykazał się skutecznością na poziomie 95.3%.

W pracy zestawiono najlepszych agentów MCTS oraz RL. Przeprowadzone eksperymenty wskazują na istotną przewagę uczenia ze wzmocnieniem. Jednak rozwiązanie to wymaga dużego nakładu obliczeniowego na etapie uczenia. Czas uczenia jest silnie zależny od dostępnej mocy obliczeniowej. MCTS ma pozornie mniejsze wymagania, jednak czas decyzji jest znacznie większy.

Dr hab. inż. Aleksandra Kawala-Sterniuk
Profesor Uczelni
Wydział Elektrotechniki, Automatyki i Informatyki
Politechnika Opolska
ul. Prószkowska 76
45-758 Opole
a.kawala-sterniuk@po.edu.pl

Opole, 18.01.2023 r.

Review Report on Doctoral Thesis

PhD candidate: **Mgr inż. Konrad Godlewski**

Thesis title: **"Monte Carlo Tree Search and Reinforcement Learning methods for multi-stage strategic card game"**.

Supervisor: **dr hab. Inż. Bartosz Sawicki**

Institution: **Warsaw University of Technology (Politechnika Warszawska)**

Doctoral Study Programme/Scientific Discipline: **Technical Information Technology and Telecommunications (Informatyka Techniczna i Telekomunikacja)**

The purpose of this review is to determine whether this doctoral dissertation meets the requirements for theses at this stage of education and whether it meets the conditions listed in the following points:

1. **Relevance of the chosen dissertation topic.**

Ad. 1.

In this doctoral dissertation, the author presented the application of machine learning methods to create agents of a card game titled "The Lord of the Rings: The Card Game" (LOTRCG). The choice of the dissertation topics is fully consistent with the discipline in which the doctoral student is going to defend his thesis - which is **Technical Information Technology and Telecommunications**. The content and content of the work are fully consistent with the topic of this work. I have no objections whatsoever to fulfilling this point.

2. **Definition and fulfilment of the dissertation objectives.**

Ad. 2.

The topic of the thesis is "Monte Carlo Tree Search and Reinforcement Learning methods for multi-stage strategic card game". In this work, the author presented the application potential of machine learning methods to create agents for a card game. The nature of the described gameplay presented in this dissertation is complex, sequential, which

Handwritten signature

means that a decision made in one phase of the game affects its subsequent phases. Random events make it impossible to combine several decision-making moments.

As the proposed game is characterised by a lot of variety, it is also a big challenge for AI agents. To overcome these, the author proposed the use of two different families of computational intelligence techniques. The first of them is the Monte Carlo Tree Search (MCTS) method, which is based on heuristic search using random sampling of game states. The win rate obtained in this way is **82.8%**, which can be considered a high result. The second of the presented and analysed techniques is Reinforcement Learning (RL), which is based on a trial and error method in which the agent is interacting with the environment. The best (in the author's opinion) set of RL agents showed a very high efficiency of **95.3%**.

The author stated that the main aim of his doctoral thesis project was verification whether AI-based methods can successfully play strategic card game (LOTRCG). Experimental results were above **95%**, so such high efficiency allow to draw conclusions about the fulfillments of the assumptions of the doctoral thesis.

3. Evaluation of the obtained results in terms of significant contribution to the scientific field.

Ad. 3.

The results of the conducted work were proven with the high efficiency (win results) of **82.8%** and **95.3%** respectively.

The doctoral student, moreover, has a fairly good (as for this stage of his scientific career) publication achievements. In 2022, he published an article entitled: "Monte Carlo tree search: A review of recent modifications and applications" in the journal Artificial Intelligence Review, for **140** points. This article has already been (according to Google Scholar) cited **32** times. The author is also co-author (leading author) of the publication entitled "Optimization of MCTS player for the Lord of the Rings: The Card Game" published in 2021 in Bulletin of the Polish Academy of Sciences: Technical Sciences (for **100** points). It is a pity that the author did not include his achievements in the thesis.

4. Evaluation of the formal aspects of the thesis.

Ad. 4.

The dissertation consists of **67** pages, including the table of contents, which consists of **54** references, and an appendix with one table. It contains **7** chapters, **20** figures, **14** tables (excluding the **one** in the appendix), **24** listings and **15** equations. In the further part of this dissertation, I will point out both the positive and negative features of the doctoral thesis in its further part.

5. Dissertation content.

Ad. 5.

As mentioned above - the thesis consists of **7** chapters. The **first** one provides introduction to the topic together with the history of games and some background about Artificial Intelligence. The **second** chapter contains more detailed theoretical background on Monte Carlo Tree Search algorithm and Reinforcement learning. The **third** chapter contains description of the LOTRCD's rules, together with some code listings. In the

Stemba

fourth chapter the author described the process of MCTS agent developments. In the **fifth** chapter the proces of the RL agent development was presented. Chapter **sixth** contains information regarding conducted experiments with a (very brief) conclusions. The last - **seventh** chapter is the thesis summary. All chapters present original, innovative scientific problems' solution.

6. Positive qualities of the thesis.

Ad. 6.

The publication of the dissertation, unusual for doctoral dissertations, made a very positive impression. This may not be a substantive comment, but I rate the aesthetics of the project very highly. The described problem has been solved. Figures' and tables' quality is proper.

Based on the dissertation content the author showed the knowledge and ability to define a scientific problem, presented ways to solve it and experimentally verified the proposed methodology. Additionally, the PhD candidate authored and co-authored few, scientific papers with an impact factor, strongly related with with his scientific topic.

Also, the proposed topic is very interesting and innovative, thus, it describes a novel implementation of the game agent.

7. Negative aspects of the thesis.

Ad. 7.

Below my critical comments:

- The work is very short, even sparing, perhaps written in accordance with the principle "the less means more", which makes it seem careless.
- The dissertation does not emphasise the author's achievements, information about his publication achievements, possible awards or participation in scientific and research projects.
- The language of work is sometimes colloquial, which in my opinion does not suit the doctoral dissertation.
- Keywords do not reflect the content of the work.
- The division into chapters is more suitable for a scientific publication than for a doctoral dissertation.
- The titles of individual chapters do not fully reflect their content.
- The first chapter provides an introduction, but lacks a solid, thorough literature review, which negatively affects the reception of the entire work.
- Equations (5.4), (5.6) and (5.7) were incorrectly formulated. Please explain/correct these mistakes.
- The sixth chapter contains conclusions, but on less than half a page, which is not befitting a doctoral dissertation.
- There is also no separate discussion in which the author could include information on possible difficulties or problems, which would greatly contribute to its quality.
- The summary describes the achievements in detail, but without information about the theses defined earlier (at the opening of the doctoral dissertation), it was difficult to assess them.

Stenik

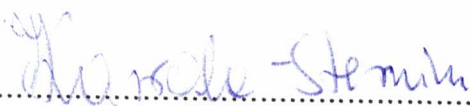
- The bibliography consists of only 54 items. This is a small number for a doctoral thesis, lower than that in many scientific publications! In addition, among the 54 bibliographic items, only 28 from the years 2017-2023 were included. With such an interesting topic - this part of the work should be more numerous.
- Information about the author's publications has not been separated in the bibliography, which is a pity.

8. Summary assessment.

Ad. 8.

Despite the short "list" of positives and the slightly longer list of negatives, the overall conclusion about the thesis is **positive**. To sum it all up, the candidate presents a thesis written in a very good English, although not always formal, which is not his first language. And in my opinion, the doctoral dissertation of **mgr inż. Konrad Godlewski** contains valuable research results and is an important scientific achievement in the development of **Technical Information Technology and Telecommunications (Informatyka Techniczna i Telekomunikacja)**, and the developed research methodology and presented results make it important for the discipline. Moreover, taking into account the author's publication achievements, I believe that this dissertation meets all the requirements for works at this stage of education and may be the subject of public defence.

I recommend the work of Mr. Konrad Godlewski for defence.


.....
Dr hab. inż. Aleksandra Kawala-Sterniuk

dr hab. inż. Jarosław Kurek, prof. SGGW
Instytut Informatyki Technicznej
Katedra Sztucznej Inteligencji
ul. Nowoursynowska 159, bud. 34, 02-776 Warszawa
jaroslaw_kurek@sggw.edu.pl

**RECENZJA ROZPRAWY DOKTORSKIEJ DLA RADY NAUKOWEJ DYSCYPLINY
INFORMATYKA TECHNICZNA I TELEKOMUNIKACJA
POLITECHNIKI WARSZAWSKIEJ**

Tytuł rozprawy: Monte Carlo Tree Search and Reinforcement Learning Methods for Multi-stage Strategic Card Game

Autor rozprawy: mgr inż. Konrad Godlewski

- 1. Jakie zagadnienie naukowe jest rozpatrzone w pracy /teza rozprawy/ i czy zostało ono dostatecznie jasno sformułowane przez autora? Jaki charakter ma rozprawa (teoretyczny, doświadczalny, inny)?**

Recenzowana rozprawa doktorska dotyczy dwóch zaproponowanych, zastosowanych i zmodyfikowanych przez autora metod uczenia maszynowego wykorzystanych do opracowania agentów dla gry karcianej "The Lord of the Rings: The Card Game" (LOTRCG). Pierwszą z nich jest Monte Carlo Tree Search (MCTS), a drugą uczenie ze wzmocnieniem (Reinforcement Learning - RL). Praca ma charakter doświadczalny tj. autor rozprawy zmodyfikował i zaimplementował wymienione algorytmy i porównał wyżej wymienione techniki inteligencji obliczeniowej.

Praca liczy 78 stron i składa się z 7 rozdziałów oraz bibliografii.

Rozdział pierwszy to wstęp, w którym autor rozprawy opisuje co jest przedmiotem niniejszej rozprawy. Przedstawiony w nim jest kontekst historyczny gier planszowych oraz wykorzystania metod sztucznej inteligencji do roli agentów w grach karcianych.

W rozdziale drugim autor rozprawy przedstawia teoretyczne aspekty dwóch rodzin algorytmów wykorzystanych w pracy, a mianowicie Monte Carlo Tree Search (MCTS) i Reinforcement Learning. Dodatkowo omówiono rozszerzenia tych metod na gry karciane.

W rozdziale trzecim przedstawiono zasady gry karcianej "The Lord of the Rings: The Card Game" obejmujące m.in. różne fazy gry z zaznaczonymi momentami decyzyjnymi. Ten rozdział zawiera również implementacje agenta strategii losowej celem analizy poszczególnych faz gry. Uzyskane przy jego pomocy wyniki pozwoliły określić kluczowe momenty decyzyjne. Ponadto agent losowy posłużył do wykonywania losowego próbkowania drzewa algorytmu MCTS.

W rozdziale czwartym znajdują się opis agenta MCTS ze wszystkimi krokami algorytmu i zastosowanymi modyfikacjami.

W rozdziale piątym autor rozprawy charakteryzuje agenta zaimplementowanego za pomocą uczenia ze wzmocnieniem z naciskiem na mechanizm kodowania/dekodowania.

W rozdziale szóstym znajdują się porównanie wyników dostarczonych przez dwóch typów zaimplementowanych agentów: za pomocą algorytmów MCTS oraz RL.

W ostatnim rozdziale siódmym autor rozprawy podsumowuje zalety i wady proponowanego rozwiązania oraz propozycje przyszłych badań.

Cel pracy został jasno sformułowany i dotyczy weryfikacji, czy agenci AI mogą z powodzeniem grać w strategiczną grę karcianą "The Lord of the Rings: The Card Game" (LOTRCG). Natomiast Teza pracy nie została jasno sformułowana co stanowi główny deficyt pracy, jednak można domyśleć się tezy pracy z wyżej wymienionego celu pracy. Praca ma znaczenie praktyczne. Podjęte badania stanowią istotny wkład w kierunku zastosowania algorytmów sztucznej inteligencji i szeroko pojętej informatyki technicznej.

W ramach głównych osiągnięć autora rozprawy można zaliczyć m.in.:

- zastosowanie i zaimplementowanie reguł eksperckich do ograniczania drzewa w algorytmie MCTS,
- zastosowanie i zaimplementowanie uczenia ze wzmacnianiem (Reinforcement Learning - RL) w częściowo obserwowalnym środowisku z różną liczbą akcji.
- zastosowanie i zaimplementowanie losowych agentów celem identyfikacji kluczowych etapów decyzyjnych w grze celem optymalizacji metryki winrate.

2. Czy w rozprawie przeprowadzono w sposób właściwy analizę źródeł / w tym literatury światowej, stanu wiedzy i zastosowań w przemyśle /świadczący o dostatecznej wiedzy autora. Czy wnioski z przeglądu źródeł sformułowano w sposób jasny i przekonujący?

W pracy wykonano dość rzetelny przegląd literatury, świadczący o rozległej wiedzy w ramach podejmowanego zagadnienia, także w odniesieniu do tematów pokrewnych. Literatura zawiera 54 pozycje, generalnie trafnie dobranych i poprawnie cytowanych. Wnioski wysnuwane są w sposób klarowny i przejrzysty.

Należy dodać że bibliografia zawiera 2 artykuły opublikowane przez autora rozprawy w następujących czasopismach:

1. Artificial Intelligence Review, 140 pkt. Impact Factor 9,588, 2022 rok
2. Bulletin of the Polish Academy of Sciences, 100 pkt., Impact Factor 1,515, 2021 rok

Szczególnie na uwagę zasługuje artykuł opublikowany w czasopiśmie Artificial Intelligence Review, który aktualnie może pochwalić się IF na poziomie 9,588 i jest to znaczące potwierdzenie dorobku pracy zawartej w pracy.

Jedna publikacja nie została uwzględniona przez autora w bibliografii i dotyczy ona materiałów pokonferencyjnych, a mianowicie:

1. Konrad Godlewski, Bartosz Sawicki, „MCTS Based Agents for Multistage Single-Player Card Game”, Proceedings of 2020 IEEE 21st International Conference on Computational Problems of Electrical Engineering, CPEE 2020, 2020, 9238707

Należy więc zauważyć, że autor rozprawy wykazuje się znaczącą aktywnością publikacyjną biorąc pod uwagę dorobek na arenie międzynarodowej w dyscyplinie informatyka techniczna i telekomunikacja (sumaryczny IF>10).

3. Czy autor rozwiązał postawione zagadnienia, czy użył właściwej do tego metody i czy przyjęte założenia są uzasadnione?

Praca generalnie poprawnie wykonana pod kątem metodologicznym. Postawiony cel pracy zostały należycie osiągnięty. Wnioski generalnie zostały poparte przemawiającymi argumentami.

4. Na czym polega oryginalność rozprawy, co stanowi samodzielny i oryginalny dorobek autora, jaka jest pozycja rozprawy w stosunku do stanu wiedzy czy poziomu techniki reprezentowanych przez literaturę światową?

Praca przedstawia oryginalne i unikatowe podejście, które zostało docenione na arenie międzynarodowej w postaci publikacji w renomowanym czasopiśmie Artificial Intelligence Review, 140 pkt. Impact Factor 9,588, 2022 rok. Zastosowanie zmodyfikowanego algorytmu Reinforcement Learning, której agent gra w strategiczną grę karcianą "The Lord of the Rings: The Card Game" (LOTRCG) nie było do tej pory przedmiotem badań. Dodatkowo zgodnie ze sztuką autor porównał oba algorytmu i wykazał znaczącą przewagę metody uczenia ze wzmocnieniem do często stosowanej metody Monte Carlo Tree Search (MCTS).

5. Czy autor wykazał umiejętność poprawnego i przekonującego przedstawienia uzyskanych przez siebie wyników /zwięzłość, jasność, poprawność redakcyjna rozprawy ?

Generalnie zastosowany aparat badawczy jest adekwatny do poruszanego zagadnienia. Wnioski wysnuwane w pracy są zrozumiałe, poparte odpowiednią analizą. Praca poprawna w zakresie edytorsko – językowym.

6. Jakie są słabe strony rozprawy i jej główne wady?

Oprócz niewątpliwych zalet sygnalizowanych w pozostałych punktach, praca w moim przekonaniu posiada także wady/niedopracowania lub fragmenty które wymagały by bardziej precyzyjnego wyjaśnienia:

1. Brak zdefiniowanej tezy pracy niewątpliwie wpływa negatywnie na odczucia recenzenta, mające przeświadczenie, że praca mogła być pisana na szybko, a przez to autor rozprawy zapomniał o tak istotnym fragmencie pracy jak teza pracy.
2. Za dużo w pracy typowych fragmentów dla pracy inżynierskiej:
 - a. diagram klas str. 22, 36,
 - b. bardziej wiele listingów kodu (np. strony: 23,24,28,29,31,32,34,35,40,41,42,43), które powinny zostać przeniesione do sekcji dodatku albo powinny być dodane do repozytorium kodu github.com podobnie jak zamieszczony symulator gry: https://github.com/kondziug/LotR_simulator,
 - c. diagram sekwencji str.34.
3. Na stronie 9 i 10 autor rozprawy opisuje działanie algorytmu MCTS, jak się ma ten opis algorytmu do klasycznego podejścia algorytmu MCTS (przeszukiwania drzewa metodą Monte Carlo). Czy algorytm MCTS zawiera fragment uczenia maszynowego (algorytm propagacji wstecznej) ? czy MCTS jest typowym algorytmem uczenia maszynowego ?
4. Na stronie 49 autor rozprawy prezentuje porównanie metod vanilla MCTS vs action reduction MCTS, czy autor rozprawy rozważał lub porównywał także ewolucyjny MCTS czy też Bridge-burning MCTS?

7. Jaka jest przydatność rozprawy dla nauk technicznych?

Praca przedstawia nowatorskie podejścia, które znacząco wpłyną na rozwój algorytmów sztucznej inteligencji, szczególnie dotyczącej budowy agentów AI w kontekście karcianych gier. Ma znaczenie praktyczne. Podjęta tematyka zasługuje na kolejne badania (kontynuacje).

8. Do której z następujących kategorii Recenzent zalicza rozprawę:

- a/ ~~nie spełniająca wymagań stawianych rozprawom doktorskim przez obowiązujące przepisy~~
- b/ ~~wymagająca wprowadzenia poprawek i ponownego recenzowania~~
- c/ spełniająca wymagania
- d/ ~~spełniająca wymagania z wyraźnym nadmiarem~~
- e/ wybitnie dobra, zasługująca na wyróżnienie

Finalnie, konkludując, stwierdzam, że recenzowana praca spełnia wymagania stawiane przez „art. 179 ust.2 Ustawy z dnia 3 lipca 2018 roku Przepisy wprowadzające ustawę – Prawo o szkolnictwie wyższym i nauce (Dz. U. z dnia 30 sierpnia 2018 r., poz. 1669) w związku art. 13 ust. 1 Ustawy z dnia 14 marca 2003 roku o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki (Dz. U. z 2003 r. nr 65, poz. 595 z późn. zm.) oraz Rozporządzenie Ministra Nauki i Szkolnictwa Wyższego z dnia 19 stycznia 2018 roku w sprawie szczegółowego trybu i warunków przeprowadzania czynności w przewodach doktorskich, postępowaniu habilitacyjnym oraz w postępowaniu o nadanie tytułu profesora” na stopień doktora i wnoszę i jej dopuszczenie do publicznej obrony.

podpis

Jaworski Kirely

Poznań, 19.01.2023 r.

dr hab. inż. Aleksandra Świetlicka
Poznan University of Technology
Institute of Automatic Control and Robotics
60-965 Poznań, ul. Piotrowo 3A

Review report

of the doctoral dissertation by Konrad Godlewski, MSc, entitled:
Monte Carlo Tree Search and Reinforcement Learning methods for multi-stage strategic card game

The following review report is prepared on the basis of the resolution of the Scientific Council of the Discipline of Information and Communication Technology of the Warsaw University of Technology, of October 18, 2022 and was made in accordance with the meaning of the Act of 20 July 2018, the Law on Higher Education and Science.

1. Thematic scope of the dissertation

The thesis presents the implementation of an agent supporting the player in decisions made during The Lord of The Rings: The Card Game. The implementation of the game itself, without improvements to the existing solutions/algorithms, is a unique challenge due to the number of available expansions of the core set, and thus the huge number of cards involved in the game and the number of possible game scenarios. The game belongs to the group of collectible card games and is a cooperative Living Card Game.

The presented solutions include two methods: Monte Carlo Tree Search (MCTS) and Reinforcement Learning (RL). The first one is based on a heuristic search that uses random sampling of game states, while the latter one is based on a trial-and-error method in which the agent interacts with the environment. The Author proposes to extend the MCTR to include action reduction, which reduces the tree size thanks to elimination of cards with low utility value. In turn, due to the operation of the agent in an environment with a variable number of actions, the selection of RL algorithms was based on the method of coding actions. In the work, the author proposed two types of coding: macroactions, which allow to obtain a fixed number of actions, and direct actions, i.e. playing individual cards by the agent.

The proposed topic of the work is exceptionally interesting/innovative for several reasons. The first is the implementation of the game agent, which has not been implemented before and which is very extensive and presents a unique challenge. On the other hand, the gaming industry, apart from being very profitable, is a source of development of new technologies in areas such as optimization, artificial intelligence, machine learning and graphics processing, and so on.

2. Content of the dissertation

The dissertation has 67 pages, consists of 7 chapters and a list of literature used in the work, which includes 54 items.

The work can be somewhat unnaturally (because not according to the chapters) divided into two main parts. Chapters 1-3.3 introduce the scope of the work and present the concepts

A. Św.

necessary to understand the solutions presented in the following chapters. In turn, from the second part of chapter 3 (from 3.4) to chapter 7, proposed solutions are shown along with simulation results and discussion.

In the first chapter (Introduction), the Author introduces the history of games along with their traditional division and the division of card games due to the mechanics of the game. In addition, this Chapter briefly describes the background of Artificial Intelligence in games. As the summary of this chapter the Author presents his goal and motivation. The goal is basically summed up in one sentence: *The thesis aims to verify whether AI agents can play successfully in the strategic card game The Lord of the Rings: The Card Game (LOTRCG).*

Second Chapter provides more detailed theoretical description of Monte Carlo Tree Search algorithm and Reinforcement learning. In Chapter 3 it is possible to read about the rules of the Lord of The Rings: The Card Game, but also Author provides here first codes of functions used further in the proposed solutions.

Chapters 5 and 6 present a full implementation of the MCTS and RL agents, while Chapter 7 presents the experiments. These 3 chapters are the Author's extremely valuable contribution to the dissertation. They not only present original solutions to the problem, but also show the Author's independent research.

3. Obtained results

The aim of the thesis was to verify whether AI agents can play successfully in the strategic card game The Lord of the Rings: The Card Game.

The Author's main achievements can be summed up as:

- action-reduction extension for MCTS method
- application of Reinforcement Learning algorithms in a partially observable environment with a varying number of actions
- methodology proposed in thesis makes use of random agents to identify the key decision stages in the game

The game consists of five decisions: planning, questing, travel, defense and attack, but there are only three crucial decisions which are made with use of MCTS or RL algorithms: planning, questing and defense. In Chapter 7 Author presents different configurations of random agent combined with MCTS or RL algorithms. Author uses correctly defined measures to show which of the proposed configurations are the best.

4. Detailed comments

There are some shortcomings in the reviewed dissertation, which can be summarized as follows:

1. An undoubted disadvantage of the work is the lack of explicitly defined theses. The dissertation is written like a scientific article, where the goal is defined, while in doctoral dissertations the emphasis is placed on theses, which are tried to be proved using the content included in the dissertation. In the summary it is possible to find detailed achievements of the Author described in the reviewed work, but these are also not theses.

A. S. W.

- Equation (5.6) seems to be written incorrectly, it seems to be an assignment, while an equation came out of it.

What is the difference between s and s_t, s_{t+1} and s' ? The description does not directly imply that, and the names entered are 'state' and 'next state', so do they designate the same states? Perhaps creating an additional list of symbols at the beginning of the thesis would clarify the situation.

Equation (5.7) (because it is not a function) seems to be incomplete, since δ is a function, the notation $\delta(\cdot)$ should appear, taking into account what variables this function depends on.

Under equation (5.4) we can find: *φ is trainable variables of the actor*; is φ a vector? Are φ, w and v weights of the neural network?

Minor:

- The structure of the thesis and the names of the chapters introduce a bit of chaos. For example, Chapter 3 is not only a description of the rules of the game, as the title suggests, but also introduces some algorithms that are used in further parts of the work.
- On page 21, there is an information about simplifying the rules of the game; line 5 under the Table 3.1: *the game model does not take advantage of the card's unique abilities described in a game text (11 on Fig. 3.1)*, but there is no such item in this Figure.
- Bibliography is a bit sloppy:
 - articles are missing DOI
 - for example position [11] and [28] are missing volume, issue and page numbers

To discuss:

- Classes QNetwork, ActorNetwork and CriticNetwork use only Dense layers, have Author thought about using other layers at this stage? There is only information about motivation of using only one layer, in my opinion quite late in the thesis (few pages after defining these classes), but no information about trying to use other layers.

5. Summary - final evaluation

The reviewed doctoral dissertation presents the author's general theoretical knowledge in a given discipline, the ability to independently conduct scientific work and is an original solution to a scientific problem.

Despite some faults, in my opinion, the reviewed thesis fulfills all requirements posed on theses aimed for obtaining PhD degree. This thesis is ready to be defended orally, in front of respective committee.



.....
dr hab inż. Aleksandra Świetlicka