

Dr hab. inż. Dariusz Mrozek, prof. PS
Katedra Informatyki Stosowanej
Politechnika Śląska w Gliwicach
ul. Akademicka 16
44-100 Gliwice

Gliwice, 28 lutego 2024

SEKRETARIAT
Rady Dyscypliny AEEITK

Wpłynęło dnia25.03.2024.....

Zarejestrowano pod nr

Podpis

RECENZJA

rozprawy doktorskiej dla
Rady Naukowej Dyscypliny
Automatyka, Elektronika, Elektrotechnika i Technologie Kosmiczne
działającej
w Akademii Górniczo-Hutniczej w Krakowie

Tytuł rozprawy: Applications of reinforcement learning methodologies to autonomous driving

Autor rozprawy: Mateusz Orłowski

1. Charakterystyka zagadnienia naukowego i charakter rozprawy

Przedstawiona przez Pana Mateusza Orłowskiego rozprawa doktorska jest poświęcona realizacji wybranych problemów autonomicznej jazdy pojazdów w ruchu drogowym, planowania zachowania pojazdu w środowisku ruchu ulicznego i autostradowego, realizacji manewrów oraz kontroli prędkości i zachowania bezpieczeństwa jazdy. Główne cele i założenia rozprawy koncentrują się wokół rozwiązania tych problemów przy pomocy technik uczenia przez wzmacnianie oraz zbadania efektywności takiej realizacji. W swoich pracach Autor przedstawił problem w postaci częściowo obserwowalnego procesu decyzyjnego Markowa (ang. *Partial Observable Markov Decision Process*, POMDP), proksymalną optymalizację polityki bazując na metodzie *PPO-Clip* i zasymulował w środowisku symulacyjnym *TrafficAIEnv* różne ww. elementy jazdy autonomicznej. Zarówno teza oraz cele pracy, jak i motywacja prowadzonych badań w tym obszarze zostały sformułowane w sposób jasny i wyczerpujący. Charakter rozprawy określiłbym jako **eksperymentalno-wdrożeniowy**, ponieważ Autor zaproponował własne podejście do procesu sterowania pojazdem oparte na metodach uczenia przez wzmacnianie dla bezkolizyjnej, komfortowej jazdy w warunkach zmiennego ruchu autostradowego, parkowania w różnych scenariuszach i współpracy wielu uczestników ruchu ulicznego w przypadkach zwięzienia trasy, wymijania i przejazdu przez skrzyżowanie. Następnie dla potwierdzenia słuszności przyjętych rozwiązań Autor przeprowadził szereg badań eksperymentalnych weryfikujących wydajność utworzonych modeli dla zadanych wzorców ruchu ulicznego i zachowania pojazdu w środowisku drogowo-parkingowym, które pomogły mu zweryfikować postawioną hipotezę badawczą.

2. Umieszczenie problemu rozpatrywanego w rozprawie w kontekście światowej literatury

Analiza światowej literatury i bieżącego stanu wiedzy w omawianym obszarze zostały przeprowadzone w sposób wystarczający i świadczą o dostatecznej wiedzy Autora w tej dziedzinie. Pokazują one ponadto, że problem podjęty w przedłożonej rozprawie jest istotny i wart dalszych badań. Szczególną uwagę zwraca Autor na wyzwania związane ze sterowaniem pojazdem w sposób zautomatyzowany w środowisku autostradowym przy pomocy interfejsu wysokiego poziomu, a także problemy uzyskania ścieżki pozwalającej na doprowadzenie samochodu do miejsca parkingowego w bezkolizyjny sposób, jednocześnie minimalizując konieczne zmiany kierunków oraz rozwiązanie problemu koordynacji ruchu wielu pojazdów jednocześnie w scenariuszach drogowych wymagających ścisłej współpracy od uczestników ruchu. Przeprowadzony przez Autora przegląd wiedzy w zakresie pokrewnych rozwiązań przedstawiony w rozdziałach 3.3, 4.3 i 5.3 pozwoliły mu w sposób jasny i przekonujący sformułować wnioski dotyczące istniejących podejść, a także określić adekwatny sposób reprezentacji obiektów środowiska, przyjąć kinematyczny model ruchu pojazdu, opracować mechanizmy detekcji kolizji z przeszkodami oraz dedykowane modele uczenia przez wzmacnianie wraz z odpowiednią funkcją nagrody.

3. Poprawność rozwiązania i przyjętych założeń

Na początku realizacji rozprawy Pan Mateusz Orłowski zdefiniował cele rozprawy, do których realizacji konsekwentnie dążył w swoich pracach badawczych. Dotyczyły one w ogólnym ujęciu zbadania możliwości automatyzacji sterowania pojazdem z wykorzystaniem uczenia przez wzmacnianie. W swoich pracach Autor oparł się na typowej architekturze systemu autonomicznej jazdy obejmującej elementy sensoryczne, percepcyjne, planistyczne i sterujące. W module planowania zachowania w środowisku autostradowym (rozdział 3) wprowadził m.in. deterministyczny zbiór zasad bezpieczeństwa i mechanizmy ograniczenia przestrzeni akcji. Wprowadził także maszynę stanów dla wykonywanych przez pojazd manewrów. W module odpowiedzialnym za znalezienie ścieżki parkowania (rozdział 4) opracował program nauczania pozwalający na stopniowe zwiększanie trudności scenariuszy parkowania. Opracował również dwa modele sieci neuronowych przyjmując różne sposoby reprezentacji sceny (wolne miejsca i kształty geometryczne). W module związanym z manewrowaniem wieloagentowym (rozdział 5) wprowadził i zweryfikował różne sposoby nagradzania, w tym nagrodę współdzieloną. Na podstawie lektury wyników eksperymentów przedstawionych w końcówkach rozdziałów 3-5 rozprawy można stwierdzić, iż postawione w rozprawie cele udało się osiągnąć, a omawiane zagadnienia zostały rozwiązane w sposób właściwy lub pozwalający sądzić o przydatności rozwiązania w zastosowaniach przemysłowych. Autor osiągnął to poprzez:

- 1) opracowanie właściwego sposobu reprezentacji pojazdów, przestrzeni i środowiska, w którym operują,
- 2) dobór odpowiednich metod należących do grupy sztucznej inteligencji i opracowanie funkcji nagrody,
- 3) badania eksperymentalne weryfikujące przydatność opracowanych rozwiązań w rozwijanym środowisku symulacyjnym. Wyniki przeprowadzonych przez Autora rozprawy badań potwierdziły, iż założenia co do możliwości realizacji każdego z trzech zadań okazały się słuszne i uzasadnione. W rozprawie Autor udowodnił m.in., iż hybrydyzacja deterministycznych reguł i algorytmów, takich jak maszyny stanu, predefiniowanie dozwolonych akcji i mechanizm generowania trajektorii z algorytmami uczenia przez wzmacnianie prowadzi do poprawy działania algorytmu oraz szybszego uczenia. Badania pozwoliły również porównać ze sobą różne modele sztucznych sieci neuronowych wytrenowanych dla

problemu parkowania pod kątem ich skuteczności i wydajności procesu inferencji. Wreszcie Autor udowodnił, iż spośród trzech testowanych eksperymentalnie sposobów nagradzania, ten oparty o współpracę agentów i współdzieloną nagrodę przyniósł najwięcej korzyści w przypadkach natężonego ruchu drogowego.

4. Oryginalność rozprawy i wartość rezultatów na tle literatury światowej

Przedstawiona rozprawa stanowi bardzo dobre uzupełnienie bieżącego stanu wiedzy światowej w zakresie automatyzacji jazdy w ruchu ulicznym i autostradowym. Pan Mateusz Orłowski zaproponował ciekawe, oparte na metodach sztucznej inteligencji podejścia w zakresie efektywnej kontroli manewrowania pojazdem zorientowanego na konkretne cele, a także przeprowadził proces ich wnikliwej oceny. Na rozprawę składa się sześć rozdziałów, które całościowo ujmują tematykę rozprawy:

Rozdział 1 zawiera ogólne wprowadzenie do tematyki pracy, motywując podjęty w trakcie realizacji tematu cel, formułuje podjęte problemy oraz artykułuje hipotezę badawczą wraz z towarzyszącymi jej twierdzeniami pomocniczymi, a także przedstawia strukturę rozprawy. Przedstawiona teza jest bardzo ogólna, ale wsparta dodatkowymi twierdzeniami, które prowadzą do uszczegółowienia celów badawczych.

Rozdział 2 formuje podstawy wiedzy niezbędnej do zrozumienia treści rozprawy. Rozpoczyna ją podrozdział poświęcony uczeniu przez wzmacnianie, po którym następuje przegląd metod i algorytmów używanych do rozwiązania różnych problemów powiązanych z dziedziną tego typu uczenia. W części tej dokonano również przeglądu prac pokrewnych wykorzystujących różne ewolucje wskazanych metod. Odrębną część rozdziału stanowi ta poświęcona jeździe autonomicznej z przedstawieniem poziomów autonomii SAE podczas realizacji jazdy wspomaganiej, ogólnej architektury systemu autonomicznej jazdy, przeglądem używanych rodzajów czujników, a także problemom lokalizacji pojazdów, ich śledzenia, fuzji danych, oceny i predykcji sytuacji na drodze, a także planowania ścieżki i trajektorii ruchu.

Rozdział 3 jest poświęcony zorientowanemu na konkretne cele (utrzymanie pasa ruchu) planowaniu zachowania pojazdu poprzez sterowanie wykonywanymi przez niego manewrami i prędkością. Przyjęto tutaj politykę interakcji ze środowiskiem polegającą na zachowaniu równowagi pomiędzy komfortem jazdy (przyspieszenie), jego bezpieczeństwem (niedopuszczenie do stanów terminalnych) i efektywnością (osiągnięcie celu jak najszybciej). W rozdziale przedstawiono hybrydowe podejście do planowania zachowania pojazdu poprzez włączenie do technik sztucznej inteligencji metod deterministycznych, bazujących na regułach. Autor opracował także algorytm planowania trajektorii ruchu oraz przeprowadził proksymalną optymalizację polityki bazując na metodzie *PPO-Clip*. Badania pozwoliły stwierdzić na jakim etapie tworzenia modelu planowania zachowania lepiej jest włączyć elementy deterministyczne (w trakcie, czy po fazie trenowania). Potwierdziły one także szybkość i skuteczność metody planowania zachowania i trajektorii jazdy w różnych scenariuszach ruchu drogowego.

Rozdział 4 zawiera opis rozwiązań dotyczących sterowania parkowaniem pojazdu w wariancie prostopadłym, równoległym i pod kątem do osi jezdni. Autor przedstawił kinematyczny model użyty

w symulacji zadania parkowania. Założył różne reprezentacje obiektów sceny w postaci grafu przeszkód oraz wolnej przestrzeni parkingowej i opracował dla nich odpowiednie architektury sztucznych sieci neuronowych. Opracował następnie taką politykę sterowania i funkcję nagrody, aby zrealizować manewr parkowania w sposób bezkolizyjny minimalizując jednocześnie zmiany kierunków. Efektywność rozwiązań została przetestowana eksperymentalnie w środowisku symulacji komputerowej oraz z wykorzystaniem rzeczywistego pojazdu dla różnych scenariuszy parkowania oraz poziomów trudności tego manewru.

Rozdział 5 rozprawy jest poświęcony manewrom ruchu drogowego, w którym zaangażowanych jest wielu uczestników. Zweryfikowano w nim zastosowanie uczenia przez wzmacnianie w trzech wybranych scenariuszach drogowych, które wymagają szerszego współdziałania uczestników ruchu. Scenariusze te przedstawiono jako problem wieloagentowy, w którym wszyscy agenci są sterowani za pomocą tej samej polityki, która decyduje o podejmowanej akcji. Podejmują oni decyzje na podstawie indywidualnego punktu widzenia, bez komunikowania się pomiędzy sobą. W części eksperymentalnej porównano wyniki dla opracowanych, różnych sposobów nagradzania, tj. przy użyciu nagrody egocentrycznej koncentrującej się na dotarciu do celu, nagrody egocentrycznej biorącej pod uwagę czas dojazdu oraz nagrody współdzielonej. Wyniki badań w środowisku symulacyjnym pokazały, iż ostatni z tych sposobów poprawia kooperację pomiędzy uczestnikami ruchu drogowego.

Rozdział 6 podsumowuje prace wykonane przez Autora w trakcie realizacji rozprawy doktorskiej i jego wkład w rozwój dyscypliny. Wskazano w nim jak poszczególne badania udowadniają stwierdzenia pomocnicze towarzyszące sformułowanej hipotezie badawczej. Zawarto konkluzje dotyczące przydatności technik uczenia przez wzmacnianie w dziedzinie jazy autonomicznej. Nakreślono także perspektywy dalszych prac.

Podjęcie i rozwiązanie problemów sterowania pojazdów autonomicznych w ruchu drogowym, polegających m.in. na planowaniu zachowania i trajektorii poruszania się pojazdów dla określonych celów jazdy, parkowania pojazdu w różnych warunkach parkingowych oraz realizacji skomplikowanych scenariuszy obejmujących wielu uczestników ruchu drogowego, oraz opracowanie dla nich odpowiednich podejść algorytmicznych (m.in. opracowanie autorskich algorytmów dla zadań generowania trajektorii ruchu pojazdu, adaptacyjnego trenowania modelu odpowiedzialnego za parkowanie, opracowanie ukierunkowanych na cele sposobów nagradzania uczestników ruchu, a także zoptymalizowane dla realizowanego zadania funkcje nagrody RL) uważam za istotne osiągnięcie Autora i zaliczam do oryginalnych wyników przedstawionych w rozprawie. Dodatkowo osadzenie tych algorytmów w rozwijanym środowisku symulacyjnym i realizację części badań z wykorzystaniem rzeczywistego pojazdu uważam za istotny element realizowanego doktoratu wdrożeniowego. Wyniki przeprowadzonych prac badawczych zostały opublikowane w artykułach naukowych m.in. w przypisanym do dyscypliny Automatyka, Elektronika, Elektrotechnika i Technologie Kosmiczne czasopiśmie *Electronics* (100 pkt. MNiSW). Dorobek ten uzupełniają artykuły opublikowane w materiałach konferencyjnych oraz patent. Potwierdza to zainteresowanie wynikami prowadzonych przez Autora prac badawczych oraz samą dziedziną.

5. Poprawność rozprawy i prezentacja wyników badań

Realizując pracę Pan Mateusz Orłowski wykazał dobre opanowanie umiejętności przedstawiania uzyskanych przez siebie wyników. Silną częścią pracy jest w mojej opinii część eksperymentalna (rozdziały 3-5), chociaż brakuje porównań z wynikami prac pokrewnych. Idee działania modeli i opisy eksperymentów (rozdziały 3-5) zostały zaprezentowane w sposób dość jasny, a lista skrótów i użytych oznaczeń i symboli na początku rozprawy pozwalają szybko przypomnieć sobie znaczenie niektórych z nich i zrozumieć sedno treści. Poza tym pierwsze rozdziały (1 i 2) pozwalają czytelnikowi dobrze odnaleźć się w tematyce rozprawy, a w kolejnych przedstawiono własne rozwiązania i wyniki przeprowadzonych prac badawczych (rozdziały 3-6). Oceny skuteczności rozwiązania dokonano z wykorzystaniem powszechnie stosowanych miar skuteczności i wydajności. Wyniki oceny skuteczności opracowanych rozwiązań zostały przeanalizowane i skomentowane w rozdziałach 3-6 przedłożonej rozprawy, pokazując w jaki sposób udowadniają one słuszność postawionej hipotezy badawczej i pozostałych stwierdzeń. Od strony redakcyjnej rozprawa jest napisana w dość dobrym stylu i czyta się ją płynnie.

6. Słabe strony rozprawy i jej główne wady

Przedstawiona rozprawa jest ciekawa i dotyczy istotnych problemów praktycznej realizacji autonomicznej jazdy. Zawiera ona najważniejsze konkluzje wyływające z przeprowadzonych prac badawczych, a część eksperymentalna pokazuje, iż opracowane rozwiązanie istotnie może służyć zamierzonym celom. Nie jest ona jednak pozbawiona pewnych uchybień lub niejasności, które krótko omawiam poniżej:

- a) Brakuje porównania wyników prac autora z wynikami innych prac pokrewnych. W pracy dokonano przecież niezłego przeglądu takich prac w kontekście różnorodności zagadnień i podejść, które zostały poruszone.
- b) W rozdz. 3.4 brakuje informacji czy były rozważane inne środowiska symulacji ruchu drogowego i dlaczego zdecydowano się skorzystać i rozszerzyć narzędzie *TrafficAI*.
- c) We wzorze 3.1 brakuje chyba wag poszczególnych składowych zachowania pojazdu. Dodatkowo, nie jest dla mnie jasne jak są one wyznaczone.
- d) Czym są podyktowane podane czasy wywołania poszczególnych bloków planowania trajektorii jazdy w rozdz. 3.5.5 (str. 52, 0,5s/0,25s/0,05s)?
- e) Jak dobrano parametry algorytmu optymalizacji PPO przedstawione w Tabeli 3.4?
- f) Jak dokładnie mierzono czas inferencji (tabela 4.5), który jest na tyle krótki, że łatwo w takim pomiarze o wystąpienie błędów pomiarowych?
- g) Nawiązując do tekstu na stronie 101, czym różni się architektura sztucznej sieci neuronowej zaprezentowana na rys. 5.5 od tej z pracy [7]? (Pada stwierdzenie, że jest „podobna”).
- h) Czy w eksperymentach z użyciem prawdziwego samochodu zrealizowano rzeczywisty manewr parkowania czy tylko zaplanowano jego możliwie najlepszą ścieżkę?

Mniej istotne uwagi:

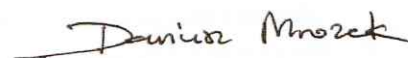
- i) Wkradło się też trochę drobnych błędów językowych (np. *tree* zamiast *three* w podpisie rys. 5.8), gramatycznych i literówek, choć nie wpływają one znacząco na odbiór treści.
- j) W pierwszej pętli *for* Algorytmu 2 brakuje wartości kończącej realizację pętli.

Uwagi o charakterze stylistycznym i językowym nie mają charakteru znacząco krytycznego i nie umniejszają znaczeniu osiągnięć Autora rozprawy. Mam natomiast nadzieję, że moje pytania staną się przyczynkiem do szerszej dyskusji, która może się wywiązać podczas obrony niniejszej rozprawy.

7. Przydatność rozprawy dla nauk inżynieryjno-technicznych

Uważam, że przedłożona rozprawa doktorska Pana Mateusza Orłowskiego wpisuje się w bieżące problemy automatyki i informatyki technicznej. Przeprowadzone prace nad realizacją wybranych elementów jazdy autonomicznej poprzez sterowanie pojazdem ukierunkowane na cele przy pomocy modeli sztucznej inteligencji i metod symulacyjnych noszą znamiona badań nowatorskich i pozwoliły Autorowi udowodnić przydatność zastosowanych podejść do omawianego problemu. W ten sposób zaproponowane rozwiązania rozszerzają spektrum istniejących prac dotyczących zautomatyzowania procesu prowadzenia pojazdów. Potwierdzają to artykuły, których Pan Mateusz Orłowski jest autorem, opublikowane przez znane wydawnictwa.

Reasumując, dobre wyniki osiągnięte przez Pana Mateusza Orłowskiego w trakcie realizowanych przez niego badań pozwalają potwierdzić główne założenia przedstawione we wstępie rozprawy. Wyniki badań pokazują, że zaproponowane przez Autora techniki i modele, a przede wszystkim ich dobra integracja ze środowiskami symulacyjnym i obliczeniowym umożliwiającymi rozpraszanie obliczeń i automatyzację wdrożenia, mogą przyczynić się do opracowania bardziej skutecznych metod automatyzacji jazdy i poprawienia jej bezpieczeństwa i komfortu w ruchu drogowym. Wartość powstałych opracowań została dostrzeżona przez środowisko naukowe, co potwierdzają opublikowane prace. Uważam zatem, że **przedstawiona rozprawa spełnia wymagania** stawiane rozprawom doktorskim określone w obowiązujących przepisach. Wnoszę o dopuszczenie Doktoranta do publicznej obrony i związanych z nią dalszych etapów postępowania.



Dr hab. inż. Dariusz Mrozek, prof. PS
Katedra Informatyki Stosowanej
Politechnika Śląska w Gliwicach