

Prof. dr hab. inż. Piotr Skrzypczyński
Politechnika Poznańska
Instytut Robotyki i Inteligencji Maszynowej

Wpłynęło dnia 20.03.2024
Zarejestrowano pod nr
Podpis *dm*

Poznań, 10.03.2024

**RECENZJA ROZPRAWY DOKTORSKIEJ DLA
RADY DYSCYPLINY AUTOMATYKA, ELEKTRONIKA,
ELEKTROTECHNIKA I TECHNOLOGIE KOSMICZNE
AKADEMII GÓRNICZO-HUTNICZEJ
IM. STANISŁAWA STASZICA W KRAKOWIE**

Tytuł rozprawy: Applications of reinforcement learning methodologies to autonomous driving

Autor rozprawy: mgr inż. Mateusz Orłowski

1. Analiza strony merytorycznej rozprawy

1.1. Obszar problemowy

Recenzowana rozprawa doktorska dotyczy zagadnień z zakresu sztucznej inteligencji oraz robotyki, w kontekście ich zastosowania do sterowania ruchem pojazdów autonomicznych. Obszar problemowy obejmuje rozwój metod planowania ruchów i działań autonomicznych pojazdów, gdy konieczne jest skuteczne radzenie sobie zarówno z problemami percepcji, jak i planowania w złożonych scenariuszach drogowych.

Z treści manuskryptu wynika, że rozprawa została przygotowana w ramach programu "Doktorat wdrożeniowy", co sprawia, że zadanie rozwiązywane przez doktoranta ma wymiar praktyczny, dotyczący rozwoju skutecznych metod autonomicznego ruchu samochodu w różnorodnych i często zmiennych warunkach drogowych. Uczenie ze wzmocnieniem jako podejście do generowania polityki planowania ruchu może pomóc w opracowaniu rozwiązań, które są zdolne do szybkiego i elastycznego reagowania na różne scenariusze drogowe, poprawiając tym samym bezpieczeństwo i efektywność pojazdów autonomicznych w praktyce. Należy jednak zauważyć, że w rozprawie nie przedstawiono odrębnego opisu wdrożenia, poprzestając na częściowej weryfikacji jednego z opracowanych rozwiązań dla rzeczywistego pojazdu (podrozdział 4.8.2). Pozostałe metody i algorytmy proponowane w rozprawie zostały zweryfikowane jedynie w symulacji, jednak widoczny jest wpływ rozwiązań firmy Aptiv, z którą doktorant współpracował na wybór i sposób implementacji środowisk symulacyjnych (TrafficAI Simulation Tool, TrafficAIEnv) oraz wybrane scenariusze sytuacji drogowych.

Za główny cel rozprawy Autor obrał badanie wykorzystania metodologii uczenia ze wzmocnieniem w kontekście planowania ruchu pojazdów w systemach ADAS (Advanced Driver Assistance System) oraz pojazdach w pełni autonomicznych (AD – Autonomous Driving), a także analizę skuteczności różnych podejść w rozwiązywaniu problemów percepcji i planowania ruchu samochodów.

Główną osią rozprawy jest zastosowanie metodologii uczenia ze wzmocnieniem. Badania skupiają się na trzech głównych eksperymentach, w których agent jest uczony różnych scenariuszy planowania ruchu pojazdów autonomicznych. Eksperymenty te obejmują planowanie zachowania pojazdu na autostradzie, samodzielne parkowanie pojazdu oraz koordynację ruchu wielu pojazdów w złożonych scenariuszach drogowych. Poprzez te eksperymenty autor analizuje skuteczność i efektywność uczenia ze wzmocnieniem w kontekście jazdy autonomicznej. Wybrane scenariusza dobrze korespondują z potrzebami automatyzacji planowania ruchu, szczególnie w systemach ADAS.

Autor rozprawy zdecydował się na rozwiązanie zagadnień omawianych w literaturze dotyczącej informatyki oraz robotyki, które jednak do chwili obecnej nie zostały rozwiązane w praktyce, szczególnie z uwzględnieniem ograniczonych zasobów obliczeniowych. Mimo wyraźnej motywacji praktycznej, tematyka rozprawy daje możliwość uzyskania nowych wyników naukowych w dziedzinie robotyki i uczenia maszynowego. Określone cele pracy i zakres prowadzonych badań uzasadniają tezę, że obszar problemowy rozprawy wpisuje się w zakres robotyki, a zgodnie z obecnym stanem prawnym, w dyscyplinie automatyka, elektronika, elektrotechnika i technologie kosmiczne.

Zdefiniowana teza badawcza rozprawy (str. 2) zakłada, że metodologia uczenia ze wzmocnieniem jest skutecznym narzędziem do rozwiązywania problemów związanych z podejmowaniem decyzji i planowaniem ruchu pojazdów autonomicznych. Twierdzenia wspierające tę tezę (interpretowane jako tezy szczegółowe) obejmują: możliwość sterowania samochodem za pomocą wysokopoziomowego interfejsu przez agenta uczącego się ze wzmocnieniem, poprawę czasu treningu i wynikowej polityki poprzez wprowadzenie deterministycznych reguł, możliwość sterowania pojazdu na niskim poziomie za pomocą bezpośredniego interfejsu planowania ścieżki przez agenta uczącego się ze wzmocnieniem, rozwiązanie problemów wieloagentowej koordynacji scenariuszy pojazdów za pomocą technik uczenia ze wzmocnieniem oraz poprawę ogólnej średniej wydajności wszystkich agentów poprzez uzależnienie nagrody indywidualnego agenta od celów innych agentów. Teza ta jest formalnie poprawna, ponieważ każde z twierdzeń jest jasno zdefiniowane i może być poddane testowi w ramach eksperymentów przeprowadzonych w rozprawie. Należy jednak zwrócić uwagę, że szczegółowe tezy numer 1 i 3 są zdefiniowane bardzo ogólnie – nie jest jasne jaki rodzaj sterowania autor ma na myśli i jakie cechy pojazdu (kinematykę, dynamikę) oraz cechy środowiska (niepewność danych, dynamika) mogą zostać uwzględnione.

Pierwszy rozdział rozprawy stanowi wstęp, zawierający zwięzłe uzasadnienie dla podjęcia tematu badawczego w kontekście rozwoju technologicznego pojazdów autonomicznych, podkreślając konieczność zastosowania zaawansowanych metod planowania ruchu w celu spełnienia wymagań przemysłowych i warunków bezpieczeństwa. Autor wskazuje, że obiecującym podejściem do rozwiązania tych problemów może być uczenie ze wzmocnieniem oraz budowa systemów hybrydowych, łączących uczenie maszynowe i systemy regułowe. Wstęp zawiera także definicję problemu naukowego, w tym proponowaną tezę, oraz bardzo zwięzły przegląd treści rozprawy. Rozdział 2 prezentuje podstawowe informacje na temat metodologii uczenia ze wzmocnieniem, w tym różne podejścia do sformułowania tego problemu oraz (dość ogólnikowo) algorytmy uczenia, a także przedstawia wybrane aspekty stanu wiedzy w zakresie samochodów autonomicznych. Trzon rozprawy stanowią rozdziały 3, 4 i 5, skupiające się na opisie proponowanych metod uczenia ze wzmocnieniem i przeprowadzonych eksperymentów. Każdy z nich obejmuje definicję problemu, przegląd literatury oraz opis badań i wyniki. Rozdziały te prezentują najważniejsze nowe koncepcje i metody stanowiące oryginalny wkład doktoranta w rozwój robotyki i sztucznej inteligencji. Rozdział 3 koncentruje się na planowaniu ruchu samochodów za pomocą interfejsu wysokiego poziomu, natomiast scenariusze parkowania z niskopoziomowym interfejsem planowania ścieżki są przedstawione w rozdziale 4. Aspekt koordynacji wielu agentów w scenariuszach miejskich jest tematem rozdziału 5. Rozprawę kończy rozdział 6, prezentujący wnioski z przeprowadzonych badań, podsumowujący oryginalne osiągnięcia Autora i wskazujący możliwe kierunki rozwoju przedstawionej koncepcji.

Praca zawiera bibliografię reprezentatywną dla tematyki, która w większości przypadków adekwatnie ilustruje omawiane zagadnienia oraz dokumentuje wkład autora w postaci publikacji jego współautorstwa i patentów.

1.2. Ocena wyników oraz stopnia ich oryginalności

Recenzowana rozprawa doktorska koncentruje się na wykorzystaniu uczenia ze wzmocnieniem w planowaniu ruchu pojazdów autonomicznych, zarówno w systemach ADAS jak i w pojazdach w pełni autonomicznych. Autor bada skuteczność różnych podejść w rozwiązywaniu problemów percepcji i planowania ruchu pojazdów, prezentując oryginalny wkład w dziedzinie robotyki i sztucznej inteligencji.

W świetle znanej literatury dotyczącej uczenia maszynowego ze wzmocnieniem w robotyce, pro-

ponowane rozwiązania można uznać za oryginalne. Autor podaje krótkie opisy stanu wiedzy w każdym z zasadniczych rozdziałów rozprawy, jednak opisy te (zawarte w podrozdziałach zatytułowanych "Prior Art") są krótkie i nie wyczerpują zagadnienia prezentacji stanu wiedzy, pomijając pewne istotne dla zagadnienia publikacje. Jedyne prezentacje stanu wiedzy w rozdziale 5, dotycząca rozwiązań wieloagentowych jest bardziej rozbudowana, jednak także tutaj brak odwołań do istotnych prac, szczególnie starszych, np. prac Petera Stone¹.

Wśród prezentowanych w rozprawie nowych koncepcji oraz rozwiązań należy wyróżnić wykazanie możliwości sterowania symulowanego samochodu autonomicznego poprzez interfejs wysokiego poziomu za pomocą metodologii uczenia ze wzmocnieniem. Ponadto, zaprezentowane badania wykazały, że wprowadzenie deterministycznych reguł podczas uczenia prowadzi do lepszej wydajności końcowej i szybszej zbieżności, co stanowi istotny krok ku efektywniejszemu uczeniu maszynowemu w kontekście autonomicznej jazdy. Integracja deterministycznych reguł wymaga jednak ostrożności ze względu na potencjalne zmniejszenie przejrzystości systemu dla agenta uczącego się ze wzmocnieniem, co sugeruje, że szczegółowa teza numer dwa sformułowana była zbyt optymistycznie. Wykazano symulacyjnie możliwość skutecznego planowania ruchu pojazdu w scenariuszach parkowania z niskopoziomym interfejsem sterowania za pomocą polityki wypracowanej przez uczenie ze wzmocnieniem. Próba integracji aplikacji parkingowej opartej na uczeniu ze wzmocnieniem w rzeczywistym systemie testowania samochodów firmy Aptiv dodatkowo podkreśla potencjał proponowanego rozwiązania w praktyce. Dość istotnym wkładem technologicznym Autora rozprawy jest także budowa narzędzi i środowisk symulacyjnych wykorzystujących rozproszone zasoby obliczeniowe i hybrydową architekturę CPU/GPU do efektywnych obliczeń.

Zauważone niedoskonałości dysertacji dotyczą kilku obszarów. Poszczególne zadania (scenariusze) uczenia rozważane w rozprawie rozwiązane zostały za pomocą algorytmu Proximal Policy Optimisation (PPO). W przypadku rozdziału 3 algorytm ten testowany był w dwóch wariantach, PPO-Clip i PPO-Penalty, natomiast w rozdziałach 4 i 5 użyty został z modyfikacjami, takimi jak curriculum learning w zadaniu parkowania lub metoda General Advantage Estimation (prawdopodobnie Autor miał na myśli Generalized Advantage Estimation²) dla scenariuszy wieloagentowych. W rozprawie nie przedstawiono jednak przekonujących argumentów przemawiających za wyborem tego konkretnego algorytmu uczenia ze wzmocnieniem, spośród wielu znanych rozwiązań. Wskazano jedynie na inspiracje rozwiązaniami zagadnień, które Autor uważa za podobne, np. w pracy [91]. Należy jednak zauważyć, że inspiracje te pochodzą spoza obszaru planowania ruchu dla pojazdów autonomicznych. W rozprawie wybrany algorytm porównywany jest w różnych wersjach i z różnymi rozszerzeniami implementacyjnymi, dotyczącymi np. percepcji, jednak nigdzie nie zaprezentowano porównania z innym algorytmem, np. DQN, wspominając jedynie taką możliwość (np. str. 51). Oczekuję przedstawienia przez doktoranta bardziej przekonującej motywacji dokonanych wyborów podczas obrony rozprawy doktorskiej.

Pewne wątpliwości można mieć także co do próby implementacji systemu automatycznego parkowania w samochodzie testowym firmy Aptiv. Integracja algorytmu parkowania z rozprawy na rzeczywistym samochodzie polegała na ocenie modułu na danych rzeczywistych, gdzie algorytm był w stanie planować ścieżki do wielu miejsc parkingowych w każdej iteracji, umożliwiając skuteczne parkowanie w większości przypadków. Jednakże z tekstu manuskryptu nie wynika, czy eksperymenty te uwzględniały działanie w czasie rzeczywistym na pokładzie pojazdu, czy tylko symulację z mapą terenu powstałą na podstawie danych rzeczywistych. Można się domyślać, że mapa rastrowa powstała na podstawie danych ze skanera laserowego, w który wyposażony był samochód widoczny na rys. 4.15, nie jest to jednak opisane w manuskrypcie. Nie jest też jasne, czy podjęto próby wykonywania manewrów zaplanowanych przez algorytm za pomocą tego samochodu. Kwestie te wymagają wyjaśnienia podczas obrony.

W poniższym zestawieniu zwracam uwagę na dostrzeżone mniej istotne niedostatki merytoryczne i zadaję pytania, na które oczekuję odpowiedzi doktoranta.

¹<https://scholar.google.com/citations?user=qnwjcfAAAAAJ&hl=en>

²<https://arxiv.org/abs/1506.02438>

1. W podrozdziale 3.5 mowa jest o algorytmie zmiany prędkości z inkreментом Δ , jednak brak jego opisu matematycznego, a działanie metody zaprezentowane jest jedynie na rysunku 3.9 o słabej jakości technicznej.
2. W podrozdziale 4.4.3 opisano algorytm wykrywania kolizji, wskazując na jedno z jego ograniczeń, jakim jest uwzględnianie tylko przeszkód w postaci wielokątów wypukłych. Rozumiem, że wybór algorytmu wykrywania przecięcia wielokątów lub odcinków i wielokątów był uzasadniony względami implementacyjnymi, warto jednak zauważyć, że istnieją efektywne obliczeniowo algorytmy realizujące to zadanie dla dowolnych wielokątów prostych (F. P. Preparata, M. I. Shamos, *Geometria obliczeniowa*. Wprowadzenie, Helion, 2003).
3. W podrozdziale 4.5.1 przedstawiono architekturę sieci neuronowej wykorzystywanej w systemie automatycznego parkowania. Zawiera ona moduł “multi-head residual attention”, który nie został dokładnie przedstawiony. Nie jest jasna jego wewnętrzna struktura.
4. Czym różnią się modele ruchu pojazdu podane w rozdziale 4 (str. 72) i w rozdziale 5 (str. 98)? Wydają się być oparte na tych samych założeniach.

Pomimo braku w recenzowanej dysertacji opisu szczegółów technicznych zrealizowanych eksperymentów, przeprowadzone badania symulacyjne umożliwiły ocenę stopnia spełnienia założeń proponowanych rozwiązań uczenia ze wzmocnieniem, co z kolei pozwoliło na osiągnięcie wystarczającego stopnia przekonania o prawdziwości tez przedstawionych w rozprawie. Wyniki przedstawione w rozprawie wystarczająco dowodzą, że główny cel podjętych badań — opracowanie skutecznych metod uczenia ze wzmocnieniem dla wybranych scenariuszy jazdy autonomicznej lub wspomagania kierowcy samochodu — został osiągnięty.

1.3. Zagadnienia dyskusyjne

Poniższe uwagi dotyczą wybranych kwestii poruszonych w rozprawie. Oczekuję komentarza doktoranta dotyczącego tych uwag.

1. W rozprawie zaproponowano modele percepcji przestrzeni wokół pojazdu oparte na przetwarzaniu opisu przeszkód w postaci wielokątów lub próbkowaniu za pomocą śledzenia promieni symulowanego sensora (prosty ray tracing). Pierwsza z tych metod prowadzi do poważnych problemów podczas tworzenia architektury sieci neuronowej, a zaproponowane rozwiązanie z siecią grafową wydaje się być mało efektywne obliczeniowo. Druga metoda jest symulacją sensora 2D w rodzaju skanera laserowego. Jej użycie powoduje jednak “epizodyczność” percepcji – system ma dostęp tylko do informacji o wolnej przestrzeni w danej chwili, zapominając natychmiast poprzednie pomiary. O ile opracowanie bardziej efektywnej architektury sieci radzącej sobie z inwariantnością zapisu wielokątów wydaje się możliwe³, lecz może nie być łatwe, o tyle niejasne są motywy rezygnacji z reprezentacji rastrowej otoczenia pojazdu w symulacji, tym bardziej, że tego rodzaju reprezentacje, także w postaci map BEV (Bird’s Eye View) są powszechnie używane w rzeczywistych pojazdach autonomicznych.
2. Ważne wyniki przedstawione w rozprawie dotyczą integracji systemu uczącego się z kodowanymi “ręcznie” regułami wprowadzonymi przez twórców tego systemu. Wyniki przedstawionych eksperymentów sugerują, że dodanie takich reguł podczas uczenia niekoniecznie prowadzi do zamierzonych efektów. Proszę o komentarz jak do tego podejścia ma się koncepcja reguł uruchamianych w przypadku niepowodzenia planowania/sterowania na podstawie wyuczonej polityki. Takie rozwiązanie zaproponowano np. w artykule M. Vitelli et al., *SafetyNet: Safe planning for real-world self-driving vehicles using machine-learned policies*, IEEE International Conference on Robotics and Automation (ICRA), Philadelphia, 2022, 897–904.

³<https://ieeexplore.ieee.org/document/9533541>

2. Analiza strony formalnej rozprawy

2.1. Ocena układu pracy i redakcji manuskryptu

Recenzowana rozprawa doktorska została przygotowana w języku angielskim. Rozprawa obejmuje w kolejności: streszczenie w języku angielskim, streszczenie w języku polskim, spis treści, wykaz ilustracji, wykaz skrótów, wykaz oznaczeń, sześć rozdziałów zasadniczych i bibliografię. Rozprawa liczy 131 ponumerowanych stron, przy czym materiał poprzedzający Wstęp numerowany jest oddzielnie (str. i – xxv, na niektórych stronach brak numeracji).

Bibliografia składa się ze 171 pozycji, w tym stron internetowych, uporządkowanych w kolejności cytowania. Spośród nich nie mniej niż 65 zostało opublikowanych w okresie ostatnich 5 lat (po roku 2018), co świadczy o dobrym osadzeniu rozprawy w aktualnym stanie badań w zakresie sztucznej inteligencji i robotyki. Cytowanych jest także kilka starszych prac, które pozwoliły Autorowi umieścić swoje badania w kontekście historii rozwoju metod uczenia ze wzmocnieniem oraz pojazdów autonomicznych. Bibliografia nie budzi zastrzeżeń od strony merytorycznej i cytowana jest właściwie. W rozprawie cytowane są 4 współautorskie prace doktoranta, opublikowane w materiałach konferencyjnych lub czasopismach, a także dwa zgłoszenia patentowe (linki do Google Patents) z jego udziałem.

Układ rozprawy jest prawidłowy, a manuskrypt jest starannie opracowany pod względem edytorskim i graficznym. Użyty w rozprawie język (angielski) i terminologia są właściwe dla poruszanej tematyki, zarówno w zakresie informatyki, jak i robotyki.

2.2. Uwagi szczegółowe

Tekst rozprawy jest zasadniczo poprawny pod względem stylistycznym. Podczas lektury manuskryptu nie zauważyłem istotnych błędów językowych. Także redakcja bibliografii jest staranna. Zauważyć można jedynie pojedyncze błędy redakcyjne, jak w [27], [102], [115] wynikające zapewne z niewłaściwego wprowadzania pozycji do systemu BibTeX. Wiele z cytowanych źródeł internetowych nie ma natomiast podanej daty, co utrudnia ocenę ich aktualności.

3. Konkluzja

Uważam, że recenzowana dysertacja Pana mgr Mateusza Orłowskiego spełnia wymogi stawiane rozprawom doktorskim przez art. 187 ustawy z dnia 20 lipca 2018 r. — *Prawo o szkolnictwie wyższym i nauce*, ponieważ zawiera oryginalną koncepcję rozwiązania istotnych zagadnień w dziedzinie robotyki. Rezultaty rozwijają istniejący stan wiedzy i mają znaczenie praktyczne, co wykazano głównie przez symulacje, a w pewnym zakresie także eksperymentalnie dla zagadnienia autonomicznego parkowania.

Uzyskane wyniki w wystarczającym stopniu dokumentują poprawność proponowanych koncepcji i opracowanych rozwiązań. Uwagi krytyczne sformułowane w treści recenzji, po części mające charakter dyskusyjny, nie umniejszają mojej pozytywnej oceny oryginalności i praktyczności przedstawionych rozwiązań. Sformułowanie problemu badawczego, zaproponowane oryginalne metody jego rozwiązania, sposób przeprowadzenia badań oraz zademonstrowana umiejętność formułowania wniosków świadczą o przygotowaniu doktoranta do pracy naukowej. Na tej podstawie wnioskuję o przyjęcie rozprawy doktorskiej mgr Mateusza Orłowskiego i o dopuszczenie do jej publicznej obrony.


Piotr Skrzypczyński