

Warszawa, 20 grudnia 2023 r.

Prof. dr hab. inż. Ewa Niewiadomska-Szynkiewicz  
Wydział Elektroniki i Technik Informacyjnych  
Politechnika Warszawska  
[ewa.szynkiewicz@pw.edu.pl](mailto:ewa.szynkiewicz@pw.edu.pl)

SEKRETARIAT  
Rady Dyscypliny AEEITK

Wpłynęło dnia 04.01.2024

Zarejestrowano pod nr .....

Podpis .....

RECENZJA ROZPRAWY DOKTORSKIEJ  
opracowanej na wniosek  
**Rady Dyscypliny Automatyka, Elektronika, Elektrotechnika i Technologie Kosmiczne  
Akademii Górniczo-Hutniczej im. Stanisława Staszica w Krakowie**

**Tytuł rozprawy:** Prediction model of an autonomous vehicle's behavior, based on Artificial Intelligence methods.

(Model predykcji zachowania się pojazdu autonomicznego oparty na metodach sztucznej inteligencji.)

**Autor rozprawy:** mgr inż. Nikodem Pankiewicz

**1. Ogólna charakterystyka rozprawy. Cel badań.**

Dynamiczny rozwój branży motoryzacyjnej obejmuje m.in. intensywne prace badawcze nad nowymi, skutecznymi metodami wspierania kierowców. Obserwuje się wzrastające zapotrzebowanie na nowoczesne, kompleksowe systemy wspomaganie decyzji, podnoszące komfort prowadzenia pojazdu, a przede wszystkim bezpieczeństwo jazdy. W przypadku samochodów efektem błędnie działającego oprogramowania jest bezpośrednie zagrożenie zdrowia i życia użytkowników. Stąd jakość i skuteczność tego typu narzędzi ma kluczowe znaczenie, szczególnie w pojazdach autonomicznych, gdzie praktycznie wszystkie decyzje są podejmowane przez komputer pokładowy. Z drugiej strony gwałtowny rozwój technologii komputerowych skutkujący ogromnym wzrostem mocy obliczeniowych współczesnych maszyn i narzędzi informatycznych, w tym bibliotek numerycznych, oraz systematyczne budowanie baz danych dają możliwość szerokiego zastosowania metod uczenia maszynowego i sztucznej inteligencji do rozwiązywania szeregu problemów, w tym związanych z podejmowaniem decyzji w trybie operacyjnym.

Przedmiotem badań udokumentowanych w rozprawie są metody, algorytmy i narzędzia informatyczne wspomagające jazdę w trybie adaptacyjnego tempomatu. Celem było opracowanie, realizacja i eksperymentalna weryfikacja rozwiązań, które mogą być zastosowane w nowoczesnych pojazdach poruszających się po drogach szybkiego ruchu, w rzeczywistym środowisku, przy założeniu konieczności podejmowania szybkich decyzji w przypadku wystąpienia potencjalnych kolizji.

Cel ten był konsekwentnie realizowany, począwszy od identyfikacji problemów występujących w procesie tworzenia informatycznych systemów wsparcia kierowców, przegląd rozwiązań dostępnych w literaturze, wskazanie na występujące w nich ograniczenia, sformułowanie zadań optymalizacji oraz modeli systemu, opis metody i stosownych algorytmów, po ich realizację w postaci modułów oprogramowania i przeprowadzenie eksperymentów badawczych na danych symulacyjnych i rzeczywistych.

Rezultatem badań jest propozycja polityki sterowania, której realizacja pozwala na osiągnięcie oczekiwanej wartości przyspieszenia w zakładanym horyzoncie czasowym. Do wyznaczania decyzji Doktorant zastosował hybrydowe podejście integrujące metody uczenia ze wzmocnieniem oraz imitacji zachowań. Zaproponował również procedury poprawy jakości danych zbieranych z rzeczywistych czujników.

Realizując zadania badawcze związane z osiągnięciem celu rozprawy Doktorant sformułował tezę, w której stwierdza (str. 3), że wykorzystanie w procesie trenowania polityk sterowania bazujących na sztucznej inteligencji danych symulowanych i rzeczywistych istotnie poprawia jakość rozwiązania w stosunku do sytuacji, gdy uczenie jest realizowane tylko na danych symulowanych. Ponadto, kombinacja podejść zakładających uczenie w trybie *off-line* i *on-line* pozwala na eliminację części ograniczeń charakterystycznych dla obu metod. Słuszność tezy została potwierdzona wynikami eksperymentów wykonanych w środowisku symulatora.

W dobie upowszechniania się usług cyfrowych w dziedzinie motoryzacji, w tym rosnących oczekiwań użytkowników oraz dążenia do wprowadzania coraz większej autonomii pojazdów, uważam podjętą w rozprawie problematykę za bardzo aktualną i istotną, zarówno z poznawczego, jak i praktycznego punktu widzenia.

## 2. Syntetyczna analiza treści rozprawy. Charakter rozprawy

Zasadnicza część rozprawy składa się z ośmiu rozdziałów.

Rozdział 1 zawiera wprowadzenie w tematykę pracy. Obejmuje przedstawienie kontekstu rozważanego zagadnienia, główne problemy badawcze, cel pracy, uzasadnienie podjętego problemu badawczego oraz jego usytuowanie na tle aktualnego stanu badań. Istotną częścią tego rozdziału jest przegląd literatury tematu. Doktorant omawiając rozwiązania proponowane przez innych badaczy zwraca szczególną uwagę na ich ograniczenia. Na tym tle prezentuje swoje podejście. Materiał zawarty w rozdziale pierwszym stanowi punkt wyjścia dla treści prezentowanych w dalszej części pracy.

Rozdział 2 jest kluczowy. Opisuje sterowany obiekt, definiuje zmienne, politykę sterowania, prezentuje modele sterowanego systemu. Formułowane jest zadanie optymalizacji w postaci maksymalizacji sumy zdyskontowanych nagród, przedstawione ograniczenia, których celem jest zapewnienie bezpieczeństwa jazdy. W sekcji 2.2. opisana jest architektura typowego systemu planowania ruchu pojazdu, w który wpisuje się rozwiązanie opracowane przez Autora.

Rozdział 3 stanowi wprowadzenie do metod sztucznej inteligencji, w szczególności tych, które Autor wykorzystał w systemie planowania ruchu pojazdów. Doktorant szczegółowo omawia techniki sztucznych sieci neuronowych, w których stosuje się uczenie ze wzmocnieniem (*Reinforcement Learning* – RL), w tym strategię zakładające uczenie *off-line* i *on-line*. Opisuje różne algorytmy optymalizacji i uczenia. Zwraca uwagę na problemy związane z adaptacją rozwiązań wykorzystujących modele sieci RL wytrenowane w środowisku symulacyjnym do warunków rzeczywistych. Prezentuje dwa podejścia proponowane w literaturze zakładające losowe zmiany parametrów symulacji oraz wykorzystanie generatywnych sieci antagonistycznych (GAN).

Rozdział 4 jest kolejnym istotnym rozdziałem pracy. W pierwszej części rozdziału opisane jest środowisko pracy adaptacyjnego tempomatu, w tym przedstawione modele matematyczne oraz przyjęte wartości liczbowe parametrów. Podane są również formuły na przyspieszenie docelowe, którego wartość zależy od wyznaczonej polityki sterowania oraz na funkcję nagród. Kolejna sekcja zawiera opis zbioru danych trenujących. Szczegółowo

omówione są zastosowane metody filtracji danych, prezentowane przykładowe próbki. Bardzo ważną sekcją jest podrozdział 4.3, w której Autor szczegółowo omawia zaproponowaną 3-modułową architekturę sztucznej sieci neuronowej do wyznaczania polityki sterowania pojazdem w trybie adaptacyjnego tempomatu. Zwraca uwagę na ograniczenia jakie napotkał podczas prowadzenia badań. Były one związane z zebraniem danych uczących oraz mocami obliczeniowymi dostępnego sprzętu komputerowego.

W rozdziale 5 omawiane są metody uczenia opracowanej architektury sieci neuronowej. Jest to rozdział, w którym Autor pokazuje sposoby eliminacji niedoskonałości danych trenujących. Proponuje metody naprawy zbiorów danych uczących wykorzystujące filtracje i optymalizację. Następnie prezentuje wyniki porównawcze modeli wyuczonych na danych oryginalnych i poprawionych. Do uczenia wykorzystał dwie metody opisane w rozdziale 3: Behavioral Cloning i MARWIL.

Wyniki eksperymentów, których celem było porównanie wariantów modeli sieci wytrenowanych za pomocą różnych algorytmów, w tym zakładających wykorzystanie metod RL w trybie *off-line* i *on-line* są omówione w rozdziale 6. Autor zdefiniował w nim również kluczowe wskaźniki, według których oceniał jakość rozwiązania. Prezentowane wyniki badań potwierdzają słuszność tezy sformułowanej w rozprawie.

Rozdział 7 jest poświęcony próbie wyjaśnienia decyzji podejmowanych przez sieć neuronową wyznaczającą politykę działania tempomatu. Autor zaproponował wykorzystanie do tego celu eksperymentu symulacyjnego i analizy statystycznej. Szczegółowo przeanalizował istotność rozważanych atrybutów, ich wpływ na podejmowane decyzje oraz wzajemną korelację niektórych atrybutów.

Podsumowanie rozprawy zawiera rozdział 8. Autor wskazuje na fakt udowodnienia tezy rozprawy. Przedstawia w nim również wnioski końcowe i perspektywę dalszych badań.

### **3. Ocena analizy źródeł i sposobu sformułowania wniosków wynikających z analizy źródeł.**

Ogółem Autor w całej rozprawie odwołuje się do 128 pozycji literatury związanych z tematyką pracy. Są to prace publikowane w renomowanych czasopismach o wysokich wskaźnikach wpływu, prace konferencyjne oraz dokumentacje narzędzi programistycznych. W przypadku tych ostatnich brakuje nieco dat oraz hiperlinków do odpowiednich stron www. Przegląd obejmuje liczne pozycje z ostatnich lat.

Doktorant analizuje literaturę głównie w rozdziale pierwszym (sekcja 1.5) oraz rozdziale 3. Na podstawie literatury (sekcja 1.5) omawia różne podejścia, metody, algorytmy i narzędzia programistyczne wspierające kierowców pojazdów. Przegląd jest zawężony do rozwiązań koncentrujących się na projektowaniu wygodnego w użyciu, gwarantującego ekonomicznie efektywną i bezpieczną jazdę systemu wspomagania kierowcy. Szczególną uwagę Autor zwraca na zastosowanie metod uczenia maszynowego, a konkretnie uczenia ze wzmocnieniem, uczenia głębokiego oraz uczenia się imitacji (naśladowania eksperta) do planowania trajektorii ruchu. Specjalną sekcję poświęca różnym realizacjom tempomatów, koncentrując się na tempomatach adaptacyjnych oraz różnych metodach wykorzystywanych do ich realizacji, takich jak: regulator liniowo-kwadratowy (LQR), sterowanie predykcyjne (MPC), klasyczne metody uczenia maszynowego, sieci splotowe i inne. Szkoda, że Autor nie przedstawił nieco dokładniejszych opisów dostępnych rozwiązań prezentowanych w literaturze, w szczególności tych, w których zastosowano techniki uczenia maszynowego, w tym szczególnie sztuczne sieci neuronowe. W pracy wspomina o ograniczeniach wymienianych metod i algorytmów, może warto było zilustrować je przytaczając wybrane wyniki testów prezentowane przez ich autorów. Wyciągane wówczas wnioski wynikające z przeglądu powiązanych prac byłyby bardziej przekonujące.

Rozdział 3 jest poświęcony metodom sztucznej inteligencji, w szczególności technikom uczenia ze wzmocnieniem, które są szczegółowo omówione. Autor prezentuje ograniczenia i mankamenty stosowanych podejść do uczenia, pokazuje sposoby poprawy.

Podsumowując, uważam, że mimo pewnych braków przedstawiona w rozprawie analiza źródeł oraz postawione na podstawie przeglądu literatury wnioski świadczą o wiedzy autora w przedmiocie rozprawy. Wnioski te zostały wykorzystane przy opracowaniu autorskiego rozwiązania.

#### **4. Analiza poprawności rozwiązania przedstawionego zadania oraz główne osiągnięcia**

Chociaż problem tworzenia narzędzi do wspomagania jazdy w trybie adaptacyjnego tempomatu przyciąga uwagę wielu badaczy i inżynierów, to jednak poszukiwanie sposobów, które pozwolą na podniesienie ich efektywności oraz zwiększenie bezpieczeństwa i komfortu jazdy jest zadaniem wciąż aktualnym. Rozprawa zawiera nowe oryginalne rezultaty. W szczególności, za najważniejsze osiągnięcia uważam następujące:

- Propozycję architektury sieci neuronowej stanowiącej kluczowy komponent zaprojektowanego planera do wyznaczania docelowego przyspieszenia pojazdu, osiąganego w określonym horyzoncie czasowym. Eksperymentalne dostrojenie parametrów sieci oraz zdefiniowanie funkcji nagrody w uczeniu ze wzmocnieniem.
- Opracowanie metody poprawy zbioru danych uczących zarejestrowanych w warunkach rzeczywistych, przez filtrowanie i optymalizację danych oraz wybór odpowiednich wskaźników do oceny rozwiązania.
- Zaproponowanie metodyki uczenia sieci ze wzmocnieniem zakładającej wstępne trenowanie na zbiorze danych rzeczywistych i douczenie otrzymanych modeli w symulatorze imitującym rzeczywiste środowisko pracy.
- Opracowanie i realizację algorytmów do wyznaczania optymalnej polityki sterowania adaptacyjnym tempomatem wykorzystujących uczenie ze wzmocnieniem i implementację wspomnianej architektury sztucznych sieci neuronowych.
- Wyniki eksperymentalne pokazujące skuteczność rozwiązania, w tym analizy porównawczej z rezultatami działania testowego kierowcy oraz klasycznym podejściem wykorzystującym uczenie ze wzmocnieniem, w przypadku, gdy do trenowania sieci wykorzystuje się jedynie zbiory danych utworzone na bazie rzeczywistych pomiarów lub tylko dane z symulatora.
- Opracowanie metody wykorzystującej eksperymenty i analizę statystyczną w celu wyjaśnienia decyzji podejmowanych przez sieć neuronową wykorzystaną do wyznaczania polityki sterowania pojazdem.

Podsumowując, Doktorant rozwiązał postawione w pracy zadania oraz udowodnił tezę rozprawy. Wykorzystując wnioski z przeglądu literatury naukowej i branżowej zaproponował politykę sterowania zakładającą zastosowanie sztucznych sieci neuronowych oraz metodę ich uczenia na danych pomiarowych zbieranych w rzeczywistym systemie oraz symulatorze. Wykazał eksperymentalnie, że takie podejście poprawia wyniki działania tempomatu w stosunku do tych, uzyskiwanych przy wykorzystaniu klasycznego uczenia ze wzmocnieniem. Cel pracy został więc osiągnięty, a Autor wykazał się wiedzą i umiejętnościami niezbędnymi do samodzielnego rozwiązywania problemów naukowo-badawczych i badawczo-rozwojowych w dyscyplinie automatyka, elektronika, elektrotechnika i technologie kosmiczne.

## 5. Analiza poprawności prezentacji wyników pracy

Recenzowana rozprawa doktorska liczy 137 stron, zawiera 66 rysunków, 14 tabel oraz bibliografię obejmującą 128 pozycji literatury. Zasadnicza treść obejmuje 115 stron, pozostałe zawierają spis treści, wykaz rysunków, słownik skrótów, stosowaną w pracy notację, załącznik z prezentacją wyników eksperymentów oraz bibliografię.

Praca jest napisana w języku angielskim. Jest ona zredagowana starannie, zarówno pod względem językowym, jak i graficznym. Występują pewne błędy językowe, gramatyczne i redakcyjne, ale nie wpływają one w sposób istotny na jakość tekstu. Pod względem językowym najsłabszą częścią jest streszczenie w języku polskim, w którym pojawiają się błędy językowe oraz kontrowersyjne, niezbyt zrozumiałe lub nieprecyzyjne sformułowania, np. „strategia planuje wartość docelowego przyspieszenia”, „strategia pozyskana przy użyciu metod uczenia”, „przewidywanego zachowania strategii”, „metodyka”, czyli zbiór zasad dotyczących sposobów wykonywania działań, jest mylona z „metodologią”, czyli nauką o metodach.

Rysunki są wykonane starannie. Czasami nazwy w opisach w głównym tekście odnoszące się do rysunków nie pokrywają się z nazwami zamieszczonymi na tych rysunkach (np. nazwy komponentów architektury rozwiązania na rysunku 2.3, str. 17 nie pokrywają się z nazwami tych modułów w tekście stanowiącym ich objaśnienie). Na niektórych wykresach opis osi nie jest czytelny (np. rys. 4.9, czy „time step” jest równy 1 sekunda?

Opracowanie rezultatów eksperymentów symulacyjnych oraz ich prezentacja w postaci tabel i rysunków nie budzą zastrzeżeń. Wyniki badań są analizowane i komentowane.

Niemniej rozprawa nie należy do pozycji, które czyta się łatwo. Uwagi krytyczne dotyczą przyjętego układu pracy i treści wybranych rozdziałów. Niektóre z rozdziałów mogłyby być nieco rozszerzone, inne skrócone. Konkretnie, główne zastrzeżenia obejmują:

- Układ pracy, w którym opracowane przez autora modele, architektury sieci, algorytmy są rozproszone po kilku rozdziałach, przeplatane rozdziałami wprowadzającymi do metod uczenia maszynowego oraz sekcjami dotyczącymi eksperymentów badawczych.
- Przyjęta notacja, gdzie ten sam symbol jest wykorzystywany do oznaczania zupełnie różnych wielkości. Nie ma jednolitej polityki wprowadzania indeksów, raz są to indeksy górne, raz dolne.
- Prezentacja formuł matematycznych. Brak szczegółowych informacji o wyprowadzeniu (lub źródle) niektórych formuł, brak objaśnień niektórych zmiennych i parametrów w rozdziałach prezentujących formuły, pomyłki w równaniach.
- Dość częste odwoływanie się do definicji zamieszczonych kilka rozdziałów dalej. Nie widzę powodu, dlaczego nie mogły być przedstawione wcześniej.
- W wielu przypadkach brak informacji czy prezentowane definicje, funkcje, modele są zaczerpnięte z literatury tematu, czy też zostały zaproponowane przez autora.

Przechodząc do bardziej szczegółowych uwag. Obecny rozdział 2 „Project Foundations” robi wrażenie nieco chaotycznego. Wydaje się, że byłby bardziej czytelny, gdyby rozpoczynał go ogólny opis zadania, systemu pomiarowego (obecny rozdział 2.3) oraz architektury typowego systemu planowania ruchu (rys. 2.3), z zaznaczeniem, w którym module są wykorzystywane metody Doktoranta. Nazwy modułów na rysunku 2.3 nie

pokrywają się z nazwami w przedstawionym poniżej opisie. Obecny rozdział 2.1 mógłby być napisany lepiej, mniej chaotycznie, bez zbędnych, nadmiarowych informacji, które dodatkowo odwołują się do odległych sekcji (str. 12). Wszystkie zmienne powinny być definiowane przed ich użyciem, np.  $d_{lon\_min}$  (2.19),  $d_{lat\_min}$  (2.20),  $s_t$  (2.18). W przypadku formuł liczenia odległości  $d_{lon\_min}$  i  $d_{lat\_min}$  należało podać pozycje literatury, w których są wyprowadzone. Nie jest oczywiste czy wzór na funkcję nagrody (2.15) jest zdefiniowany przez Autora, czy przyjęty z literatury. Wskazane byłoby podanie ewentualnego źródła lub dokładniejsze objaśnienie. Warto było zakończyć ten rozdział sekcją podsumowującą, w której byłby dokładnie zdefiniowany rozwiązywany problem (najlepiej w postaci formalnej). Część tych informacji znajduje się na stronach 15 i 16. Generalnie w rozdziale brakuje odwołań do literatury.

Podobne uwagi są do rozdziału 3 „Theoretical introduction to learning algorithms”. Nie jest do końca zrozumiałe, co w tytule rozdziału ma oznaczać „Theoretical”? Sekcje poświęcone metodom przeplatają się z sekcjami definiującymi wykorzystywane funkcje. Może lepiej było zrobić jeden podrozdział definiujący używane dalej funkcje i polityki w sekcji 3.1 (czyli podrozdziały 3.1.4, 3.1.5, 3.1.6, 3.1.9), a potem kolejno omówić metody. W obecnej wersji rozdziału czytelnik może się nieco zagubić, szczególnie, że dość często występują odwołania do odległych, kolejnych rozdziałów. W tym rozdziale również, w wielu miejscach brakuje referencji do stosownej literatury, np. pozycji, z których pochodzą przedstawione definicje funkcji  $V$  i  $Q$ . Dlaczego najpierw są opisane konkretne algorytmy RL (np. rozdz. 3.1.7, 3.1.10), a później mamy klasyfikację algorytmów RL (rozdz. 3.1.12)?

Rozdział 4 „Experimental setup” znowu zawiera informacje dotyczące konkretnych wartości parametrów wykorzystywanych w eksperymentach symulacyjnych, opis zbioru danych uczących, architektury sieci neuronowych, jak też modele matematyczne otoczenia oraz definicję funkcji nagrody. Tytuł rozdziału wskazuje, że raczej należało się w nim skoncentrować na opisie założeń do eksperymentów i opisie scenariuszy testów. Modele lepiej było zamieścić wspólnie z opisami zawartymi w rozdziale. 2, a zaproponowanym architekturom sieci, łącznie z definicjami komponentów funkcji nagrody oraz dyskusją rozwiązania poświęcić oddzielny rozdział.

Rozdział 5, tabela 5.1. Niektóre parametry nie są opisane (np. KL coefficient). Nie ma też stosownych odwołań do wzorów, z których mogły być liczone. Czy parametr „clip” jest liczony zgodnie z formułą (3.29)? Str. 68, zadanie minimalizacji (5.1) – powinno być „denotes cost terms defined in Eq. 5.4, 5.5, 5.7.”

Czy rozdziały 6.2 i 6.3 są poświęcone omówieniu wyników tych samych eksperymentów? Nie jest jasne, czy są tu prezentowane wyniki uczenia czy testowania rozwiązań? W rozdziale 6 brakuje informacji czy prezentowane wyniki to rezultaty uzyskane na zbiorach testowych. o wielkościach zbiorów uczącego i testowego. Czy tabele 6.1 i 6.2 prezentują wyniki uczenia? Czy uczono sieci na tych samych danych?

Uwagi krytyczne dotyczą również prezentacji formuł matematycznych. Bardzo pomocna jest sekcja z przyjętą notacją oraz używanymi skrótami nazw metod i systemów, ale wskazane byłoby również objaśnianie używanych zmiennych i parametrów, nawet jeśli wydaje się, że nazwy są samoobjaśniające się, bezpośrednio po prezentacji zawierających je formuł, przynajmniej, gdy po raz pierwszy pojawiają się w danym rozdziale. W niektórych przypadkach takich objaśnień brakuje, np. wzory (2.1), (2.2), (2.14), (2.20), (3.17), (3.22), (5.14) itd. Nie jest również dobrą praktyką odwoływanie się do definicji lub formuł przedstawianych w kolejnych rozdziałach (np. (2.16), str. 15). Występują błędy w symbolach, np. str. 15 (brak indeksu „s”). Występują błędy we wzorach, np. w (2.16) – niejednoznaczność w przypadku  $\Delta s = d_{lon\_min}$  (chyba, że jest to akceptowalne, ale wymaga

wyjaśnienia). W formułach (3.14) i (3.15) mamy  $r_{t+1}$ , a w algorytmach 1 i 2  $r_t$ , podobnie z  $u_t$  (wzór 3.15 i algorytm 2). Niektóre zmienne są używane do oznaczania różnych wielkości:  $\tau$ ,  $s$  (raz jest wektorem stanu, innym razem indeksem),  $c$  (współczynnik aerodynamiczny, funkcja nagrody),  $x$  (współrzędna położenia, wektor parametrów) itd. Niejednolite zapisy - raz dana funkcja zależy od dwóch zmiennych, w kolejnym wzorze tylko od jednej, np. formuły (3.22) i (3.23). Algorytm 4, a w szczególności występujące w nim parametry wymagają objaśnienia. Podobnie jest z indeksami. Nie jest oczywiste, dlaczego raz są górne, raz dolne (np. mamy  $v_{max}$ , ale już  $v^{vcs}$ ). Na str. 73 autor pisze, że przyjmuje te same wagi w równaniach (5.1) i (4.6), tymczasem w jednym przypadku różnią się znakiem. Czy to jest poprawne?

Drobne zastrzeżenia można zgłosić do opracowania bibliografii – wskazana byłaby większa staranność. W przypadku niektórych pozycji literatury brakuje roku wydania, wydawnictwa, numerów stron lub nie jest wiadomo czym jest dane opracowanie (np. [92]).

## 6. Słabe strony rozprawy i jej główne wady

Uwagi krytyczne dotyczące sposobu prezentacji niektórych treści, w tym układu pracy zostały już szczegółowo przedstawione w poprzednim punkcie recenzji. Podsumowując, praca byłaby znacznie łatwiejsza w studiowaniu, gdyby Autor, po krótkim wprowadzeniu kreślącym temat i zakres pracy, przedstawił przegląd literatury, w tym opis zastosowanych metod uczenia maszynowego (część rozdziału 3), sformułował pełny model sterowanego systemu (łącznie z modelem otoczenia) i dyskusję problemu, przedstawił metodę rozwiązania problemu, w tym zaproponowane architektury sieci i algorytmy, a następnie opisał eksperymenty i przeanalizował wyniki. Szczególnie dotkliwy jest brak pełnego opisu problemu w jednym miejscu oraz brak zwięzłego opisu rozwiązania obejmującego zaproponowane architektury sieci, algorytmy oraz metodyki postępowania przy uczeniu sieci uczenia. W przedłożonej do recenzji pracy te informacje trzeba wyłuskiwać z różnych rozdziałów. Należało poświęcić więcej uwagi i bardziej uwypuklić osiągnięcia autorskie, oddzielić je od ogólnych opisów metod.

Ponadto, lektura pracy skłania do sformułowania następujących komentarzy lub uwag krytycznych:

- Co Autor rozumie przez „suboptimality of human actions” (str. 63)? Nie znajduję dokładnego wyjaśnienia w rozprawie. W czym to suboptymalne zachowanie się przejawia?
- Opracowane rozwiązanie powinno mieć charakter uniwersalny. Nie jest jasne, dlaczego autor w modelu ograniczył się do rozpatrywania konkretnie 10 pojazdów znajdujących się w otoczeniu sterowanego.
- Autor proponuje podejście pozwalające na poprawienie zbioru danych uczących w dwóch oddzielnych sekcjach. W obu przypadkach mówi o tworzeniu zbioru  $D_{opt}$ . W sekcji 5.2 pokazuje zastosowanie optymalizacji, m.in. do eliminacji zakłóceń pomiarowych czy błędnych zachowań człowieka. W sekcji 5.3 proponuje pewne reguły filtracji. Jak więc wygląda ostatecznie zbiór  $D_{opt}$  i dlaczego oba rozwiązania nie zostały opisane w jednej sekcji?
- Niektóre przedstawione formuły nie są oczywiste, np. definicja jednej ze składowych funkcji nagrody (4.2). Po pierwsze, przy założonym warunku, można spokojnie przekształcić równania i przedstawić bez operatora modułu. Po drugie, dlaczego w przypadku przekroczenia prędkości maksymalnej w równaniu nagrody mamy konkretnie liczbę 3? Generalnie przyjmowanie narzuconych, stałych parametrów, co

ma miejsce w wielu rozdziałach rozprawy, powinno być wyjaśnione i uzasadnione. Takie podejście nie jest dobrą praktyką i należy tego unikać.

- W przypadku wielu, niektórych dość istotnych, formuł matematycznych brak jest informacji czy jest to propozycja autora, czy też zaczerpnięta z literatury (nie ma podanych źródeł). Np. równania (2.19), (2.20), (5.14) i inne.
- Autor wyjaśnia, że wagi przypisane składowym funkcji nagrody (równanie 4.6) zostały wyznaczone eksperymentalnie, ale o tym jakie dane wykorzystano dowiadujemy się w następnych rozdziałach. Warto było również wspomnieć ile w tym celu wykonano eksperymentów, co pokazałoby koszt strojenia parametrów funkcji („multiple experiments” niewiele mówi).
- W kilku, zdefiniowanych na stronie 88, miarach występuje dzielenie przez zbiór  $X$ . Chyba chodziło o moc zbioru? Poza tym zapis  $X=\{t \in 1,2,\dots,T: \dots\}$  jest niepoprawny. Powinno być  $X=\{t \in \{1, 2,\dots,T\}: \dots\}$ .
- W dwóch rozdziałach mamy definicję funkcji nagrody i jej składników, tj. rozdz. 4.1.4, równania (4.2) - (4.6) i rozdz. 5.2.1, równania (5.3) – (5.13). Nazwy funkcji niby takie same, ale zależą od różnych zmiennych. Nie ma żadnych odniesień w rozdz. 5.2.1 do definicji z rozdz. 4.1.4. Na początku rozdziału 5.2.1 dowiadujemy się, że planowana trajektoria zależy od wektora  $x$  tajemniczych  $n$  parametrów, względem których jest optymalizowana. Podobnie jest z definicją polityki. Jak mają się do siebie formuły (2.13), str. 15 i (3.18), str. 31. W rozdziale 2 polityka  $\pi$  jest funkcją  $u$  (zależnego od wektora parametrów  $\theta$ ) i bieżących obserwacji, w rozdziale 3  $\pi$  jest funkcją  $u$  (zależnego od wektora stanu) i wektora parametrów  $\theta$ . O wektorze parametrów dowiadujemy się na str. 14 jedynie, że zależy on od aktualnego pomiaru stanu. Należałoby wspomnieć w tym miejscu o tym jaka jest rola tych parametrów i co one reprezentują. Na str. 58 dowiadujemy się, że są to strojone wagi neuronów.
- Rozdział 3.3 – nie bardzo rozumiem, dlaczego Autor koncentruje się na systemach robotycznych? Nie są one rozważane w pracy.
- Rysunek 4.6 jest zredagowany starannie, ale warto było dokładniej opisać zaproponowaną architekturę sieci, a przede wszystkim pokazać przepływy między modułami (co jest wejściem, a co wyjściem). Brakuje również wpisania modelu sieci w strukturę systemu planowania i utrzymywania trajektorii przedstawionego na rys. 2.3 (w którym module sieć jest wykorzystywana?).
- Nie jest oczywisty wkład Autora w tworzenie środowiska do symulacji. Czy zadaniem było tylko proste wykorzystanie dostępnych platform, czy też przeprowadzenie eksperymentów wymagało większego wysiłku, np. przygotowania niezbędnych modułów oprogramowania, dostosowania środowiska do potrzeb pracy itd. Generalnie, Autor nie zawsze dostatecznie precyzyjnie wskazuje na to, co stanowi jego oryginalne osiągnięcie.

## 7. Przydatność rozprawy dla nauk technicznych

Mimo wymienionych powyżej słabych stron pracy uważam, że przedstawione w rozprawie wyniki wnoszą istotny wkład w dyscyplinę automatyka, elektronika, elektrotechnika i technologie kosmiczne, a konkretnie w rozwój badań w zakresie systemów wspomagania kierowcy korzystającego z adaptacyjnego tempomatu. Zaproponowana metodyka uczenia sieci zakładająca wstępne wytrenowanie na rzeczywistych, historycznych danych, a następnie douczenie sieci z wykorzystaniem narzędzi symulacyjnych może mieć znacznie szersze zastosowanie.

## 8. Podsumowanie i wniosek końcowy

Uważam, że rozprawa mgr inż. Nikodema Pankiewicza w pełni spełnia wymagania Ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce (Dz. U. 2023, poz. 742) w odniesieniu do rozpraw doktorskich. W związku z tym wnoszę o przyjęcie rozprawy i dopuszczenie Autora do dalszych, przewidzianych przepisami, etapów przewodu doktorskiego.

podpis

*Nikodem*