

Warszawa, 3 grudnia 2025 r.

dr hab. inż. Paweł Wawrzyński
Instytut Badawczy IDEAS
ul. Krakowski Przemiescie 13
00-071 Warszawa

RECENZJA

rozprawy doktorskiej mgr Michała Zawalskiego

pt. *Solving Complex Decision Making Problems with Structured Reasoning and Planning*

Formalną podstawą opracowania recenzji jest uchwała Rady Dyscypliny Naukowej Informatyka Uniwersytetu Warszawskiego z dnia 25 września 2025 roku. Oceny rozprawy doktorskiej dokonano według kryteriów określonych w ustawie z 20 lipca 2018 r. *Prawo o szkolnictwie wyższym i nauce* oraz *Uchwały 157 Senatu UW z dnia 29 czerwca 2022 r. w sprawie określenia sposobu postępowania w sprawie nadania stopnia doktora oraz stopnia doktora habilitowanego na Uniwersytecie Warszawskim*. Promotorem rozprawy doktorskiej jest dr. hab. Piotr Miłoś prof. UW.

Charakterystyka rozprawy

Rozprawę stanowi zbiór pięciu artykułów:

- Artykuł P1 został opublikowany na konferencji AAMAS (CORE A*, 200 pkt. MNiSW) w roku 2022. Kandydat jest pierwszym spośród czterech jego autorów.
- Artykuł P2 został opublikowany na konferencji NeurIPS (CORE A*, 200 pkt. MNiSW) w roku 2021. Kandydat jest trzecim spośród ośmiu autorów.
- Artykuł P3 został opublikowany na konferencji ICLR (CORE A*, 200 pkt. MNiSW) w roku 2023. Kandydat jest pierwszym z dziewięciu autorów.
- Artykuł P4 został opublikowany na konferencji CoRL (niesklasyfikowanej, ale już renomowanej w obszarze uczenia maszynowego) w roku 2024. Kandydat jest pierwszym z sześciu autorów.
- Artykuł P5 został opublikowany w czasopiśmie Transactions on Machine Learning Research (cieszy się ono renomą, którą jednak przeocza MNiSW). Kandydat jest pierwszym z ośmiu autorów.

Na podstawie tego zestawienia można stwierdzić, że kandydat odgrywała pierwszoplanową rolę w szeregu prac badawczych, których wyniki zostały opublikowane na najbardziej renomowanych konferencjach, na których prezentowane są wyniki badań w dziedzinie uczenia maszynowego.

Sam autoreferat oceniam jako napisany dosyć niechlujnie. Autor podaje oznaczenia artykułów ze zbioru jako P1 – P5, by potem używać tych oznaczeń tylko w sekcji „Introduction” autoreferatu. Artykuł P1 w ogóle nie jest w autoreferacie omówiony. Pozostałe prace zostały omówione nazbyt powierzchownie.

Przedmiot rozprawy

Przedmiotem pracy P1 formułuje wersję wieloagentową znanego algorytmu uczenia ze wzmocnieniem Aktor-Krytyk z Powtarzaniem Doświadczenia i dowodzi zbieżności tej wersji pod pewnymi warunkami. Artykuł jest poza zakreślonym w rozprawie nurtem tematycznym, natomiast nie jest od niego daleko, ponieważ dotyczy podejmowania decyzji w dynamicznym środowisku w warunkach niepewności.

Przedmiotem pracy P2 jest metoda kSubK stanowiąca generator korzystnych stanów pośrednich na potrzeby algorytmów przeszukiwania przestrzeni stanów. Generator ten wskazuje stany w odległości określonej przez hiperparametr. Metody takie jak Best First Search (BestFS) i Monte-Carlo Tree Search (MCTS) wsparte przez kSubK osiągają większą skuteczność i szybkość w znajdowaniu wyników (choć za cenę wcześniejszego wyuczenia kSubK na rozwiązaniach innych problemów tego typu).

Praca P3 stanowi istotne rozwinięcie P2 i przedstawia kompleksowy algorytm przeszukiwania przestrzeni stanów AdaSubS, w którym działa szereg uczących się komponentów: generator stanów pośrednich w różnej odległości od bieżącego, polityka docierania do zadanego stanu pośredniego, weryfikator oceniający, czy dany stan pośredni jest faktycznie osiągalny z bieżącego i funkcja wartości oceniająca długość drogi od zadanego stanu do rozwiązania. Algorytm ten poprawia wyniki uzyskiwane z użyciem kSubK.

Praca P5 stanowi obszerne studium porównujące różne podejścia do przeszukiwania przestrzeni stanów, z którego wynikają przesłanki wskazujące, że metody z hierarchią będą skuteczniejsze i szybsze niż metody płaskie. Te przesłanki to: trudna do wyznaczenia funkcja kosztu dotarcia do celu, duża przestrzeń akcji, obecność ślepych zaułków w przestrzeni stanów oraz dostępność danych treningowych z rozwiązań różnych problemów.

W pracy P4 autorzy budują system sterowania robotem oparty o wizyjny model językowy. System ten otrzymuje definicję zadania robotycznego i obraz sceny. Generuje jednokierunkowy tok rozumowania (ang. chain of thought), który prowadzi go do planu działań i konkretnych akcji sterujących robotem. Autorzy realizują taki system z użyciem ciężkiego wizyjnego modelu językowego Gemini 1.0. Ten system generuje dane uczące dla systemu typu VLA (ang. vision-language-action) opartego o lżejszy wizyjny model językowy. Następuje tu więc dystylacja wiedzy.

W mojej ocenie prace przedstawione w rozprawie są bardzo dobre, ich zaprezentowanie na najbardziej prestiżowych konferencjach w branży jest w pełni uzasadnione. Stanowią one bardzo mocny materiał na wysokiej jakości rozprawę doktorską. Lektura tych prac sprawiła mi dużą przyjemność, a jej zakończenie przyjąłem z żalem.

Spójność rozprawy

Rozprawa stanowi dzieło naukowe poświęcone rozwojowi metod podejmowania decyzji w dynamicznym środowisku opartych o planowanie i uczenie się. Od tego nurtu nieco odbiega praca P1, w której wątek planowania jest nieobecny. To odstępstwo traktuję jednak jako niewielkie, a całą rozprawę uważam za spójne dzieło naukowe.

Pytania i wątpliwości

Każda dobra praca naukowa jest punktem wyjścia do postawienia kolejnych pytań. Każda praca musi być dostosowana do formatu wydawnictwa, zatem musi pozostawić pewne pytania bez odpowiedzi. Prace przedstawione w cyklu nie stanowią w tym zakresie wyjątku. Tym niemniej, uważam, że prace dostarczyły odpowiedzi na postawione w nich pytania i generalnie nie wymagają dodatkowych wyjaśnień. Jedno zagadnienie przytoczę tutaj jako wyjątek, ponieważ uznałem je za szczególnie interesujące.

W systemie przedstawionym w pracy P4 mamy system uczący się na podstawie rozwiązań wygenerowanych przez model Gemini 1.0. Można sobie wyobrazić system na bazie Gemini 1.0, który

realizuje to samo zadanie sterowania bez żadnego uczenia. Czy uczenie mniejszego modelu osiągało jeszcze jakiś cel poza ograniczeniem mocy obliczeniowej angażowanej w trakcie sterowania?

Wnioski końcowe

Rozprawa doktorska mgra Michała Zawalskiego stanowi oryginalne rozwiązanie problemu naukowego i wskazuje na wysoki poziom wiedzy teoretycznej Autora w dyscyplinie informatyka, a także na umiejętność samodzielnego prowadzenia przez niego pracy naukowej.

Stwierdzam, że opiniowana rozprawa doktorska spełnia wymogi ustawy z dnia 20 lipca 2018 r. *Prawo o szkolnictwie wyższym i nauce*. Oceniam ją pozytywnie. Wnoszę o dopuszczenie mgra Michała Zawalskiego do publicznej obrony pracy doktorskiej. Ponadto stwierdzam, że poziom naukowy rozprawy jest bardzo wysoki i wnioskuję o jej wyróżnienie.

Elektronicznie podpisany przez
PAWEŁ WAWRZYŃSKI
2025.12.04 19:02:08 +01:00