

RECENZJA ROZPRAWY DOKTORSKIEJ DLA
RADY NAUKOWEJ INSTYTUTU PODSTAW INFORMATYKI PAN
W DYSCYPLINIE INFORMATYKA TECHNICZNA I TELEKOMUNIKACJA

Autor recenzji: prof. dr hab. Andrzej Szałas

Wydział Matematyki, Informatyki i Mechaniki
Uniwersytet Warszawski

31 marca 2026

Autor rozprawy: mgr inż. Dominik Kurpiewski

Tytuł rozprawy: *E-voting, Card Games, and Drone Teams: Verification of Strategic Properties in Multi-Agent Systems with Imperfect Information*

1 WPROWADZENIE

Rozprawa dotyczy problemu formalnej weryfikacji zdolności strategicznych w systemach wieloagentowych działających w warunkach niedoskonałej informacji. Autor proponuje kompleksowe podejście oparte na aproksymacyjnej weryfikacji własności wyrażanych w logice ATL_{ir} , wykorzystujące rachunki stałopunktowe dla wyznaczania dolnych i górnych ograniczeń wartości formuł. Kluczowym elementem jest translacja do epistemicznego μ -rachunku oraz wprowadzenie nowych technik redukcji złożoności, w tym zoptymalizowanych struktur danych, abstrakcji modelu i redukcji częściowo-porzadkowych. Dla systemów asynchronicznych zaproponowano metodę lokalnej aproksymacji.

Dodatkowo przedstawiono algorytm syntezy strategii oparty na przeszukiwaniu z relacją dominacji, rozszerzony na przypadek wielokryterialny. Proponowane metody zostały zaimplementowane w autorskim narzędziu STV i przetestowane na szeregu złożonych przypadków praktycznych. Wyniki eksperymentalne wskazują na znaczącą poprawę wydajności względem istniejących rozwiązań, umożliwiając analizę modeli o dużej przestrzeni stanów.

Tematyka rozprawy plasuje się w ważnej dziedzinie weryfikacji modelowej systemów asynchronicznych, w tym z niepełną informacją. Badania tego nurtu mają duże znaczenie praktyczne i stawiają niebanalne pytania natury teoretycznej.

2 ZAWARTOŚĆ ROZPRAWY

Rozdział 1

Pierwszy rozdział stanowi wstęp do rozprawy, jest poświęcony ogólnemu przedstawieniu problemu weryfikacji modelowej, jej praktycznych aspektów, wyzwani

przed nią stojących, oraz dziedzin zastosowań. Autor prezentuje także główną tematykę rozprawy, zgodnie z którą zaprezentowane badania dotyczą algorytmów weryfikacji dla logiki ATL w tym z ograniczoną informacją (*imperfect information*), ze szczególnym uwzględnieniem praktycznych zastosowań, m.in. systemów głosowania elektronicznego. Przedstawioną tu tezę rozprawy można streścić jako stwierdzenie, że modelowanie systemów wieloagentowych jako struktur gier synchronicznych i asynchronicznych pozwala lepiej specyfikować zachowania strategiczne agentów oraz skuteczniej weryfikować ich właściwości w scenariuszach z ograniczoną informacją.

Końcowa część wstępu prezentuje dorobek publikacyjny Autora oraz jego wkład w poszczególne opublikowane wyniki wchodzące w skład rozprawy. Dorobek stanowiący podstawę rozprawy obejmuje 11 artykułów, w tym 9 na konferencjach oraz 2 w czasopismach. Warto podkreślić, że wśród artykułów konferencyjnych są 4 na konferencjach rangi A* (jeden na IJCAI, trzy na AAMAS), a jeden artykuł czasopismowy jest opublikowany w *Artificial Intelligence* będącym czołowym światowym czasopismem z dziedziny sztucznej inteligencji.

Dorobek publikacyjny nie wchodzący w skład rozprawy obejmuje 14 artykułów, w tym kilka opublikowanych na konferencjach rangi A*. Można więc stwierdzić, że dorobek ten jest bardzo bogaty.

Rozdział 2

Prezentacja rozdziału opiera się na pracach Jamroga i in. (2019a, 2020b, 2022a). Rozdział wprowadza do teoretycznych podstaw badań, tworząc ramy dla dalszych części pracy. Autor rozpoczyna od formalnej definicji problemu weryfikacji modelowej, która stanowi punkt wyjścia dla przedstawionego cyklu badań. Wprowadzono w nim podstawowe pojęcia i formalizmy, na których opierają się dalsze części rozprawy, w tym struktury gier z ograniczoną informacją (*imperfect information concurrent game structure*), warianty logiki ATL, w tym z operatorami dotyczącymi strategicznych możliwości grup agentów oraz z operatorem wiedzy, modele asynchronicznej interakcji, a także epistemiczny rachunek μ . Następnie omówiono strategię, złożoność wariantów ATL oraz znane podejścia i dostępne narzędzia weryfikacji.

Rozdział zawiera również aspekt praktyczny, prezentując przykłady narzędzi ułatwiających definiowanie i analizę modeli, co pozwala czytelnikowi lepiej zrozumieć omawiane techniki. Poprzez połączenie podstaw teoretycznych z praktycznymi zastosowaniami, rozdział zapewnia przystępne wprowadzenie do kluczowych pojęć leżących u podstaw badań.

Warto zauważyć, że ograniczona informacja jest modelowana poprzez relacje równoważności na stanach, odzwierciedlające nieodróżnialność stanów przez agentów. Takie podejście do nieodróżnialności jest charakterystyczne dla zbiorów przybliżonych (*rough sets*), gdzie relacje równoważności prowadzą do dolnych i górnych

aproxymacji. Zaskakujące jest dla mnie całkowite pominięcie tego wątku: dziedzina zbiorów przybliżonych jest bardzo dobrze rozwinięta i zbadana, a rozważane tam aproxymacje z pewnością są warte porównań z technikami proponowanymi w rozprawie.

Rozdział 3

Ten rozdział, bazujący na pracy Jamroga et al. (2019a), jest zasadniczy dla proponowanego podejścia.

Weryfikacja zdolności strategicznych w warunkach niepełnej informacji jest złożonościowo trudna. W rozdziale jest więc badana alternatywna metoda zmniejszania złożoności weryfikacji – zamiast dokładnej weryfikacji modelowej stosuje się przybliżenia w postaci dolnych i górnych ograniczeń formuł przez ich translacje kodujące warunki wystarczające (tr_L) i konieczne (tr_U). W ten sposób uzyskuje się dolną i górną aproxymację wejściowej formuły. Takie podejście opiera się na właściwościach definiowalnych punktami stałymi i mapowaniu formuł ATL_{ir} na odpowiednią odmianę alternującego rachunku μ , co zmniejsza złożoność obliczeń. Zaproponowane podejście modyfikuje semantykę operatora następnego momentu czasu w celu uzyskania ogólnej definicji ograniczenia dolnego.

Bardzo interesujące jest twierdzenie 3.1.5, które pokazuje że translacja tr_{L3} daje w wyniku poprawną aproxymację dolną dla dowolnej formuły ATL_{ir} . Badane jest także zagadnienie kiedy uzyskane aproxymacje są ściśle (*tight, exact*). Choć intuicyjnie jasne, nigdzie nie znalazłem dokładnej definicji rozumienia *tight* oraz *exact* w badanym kontekście.¹

W aspekcie uzyskiwania najściślejszych aproxymacji jest w literaturze znane podejście do badania najslabszych warunków wystarczających (*wsc*) i najsilniejszych warunków koniecznych (*snc*), wprowadzone w Lin, F. 2000. *On strongest necessary and weakest sufficient conditions*. Proc. 7th KR'2000, 167–175. Warunki *wsc* i *snc*, będące dolną i górną aproxymacją w sensie definicji przyjętych w rozprawie, były badane w rozmaitych logikach, również w kontekście zbiorów przybliżonych. Wprawdzie w swojej istocie wymagają one egzystencjalnego lub uniwersalnego fragmentu logiki II-go rzędu, ale dla wielu klas formuł kwantyfikatory te można eliminować uzyskując w wyniku formuły I-go rzędu lub stałopunktowe – por. Doherty, P.; Łukaszewicz, W.; and Szalas, A. 2001. *Computing strongest necessary and weakest sufficient conditions of first-order formulas*. Proc. 17th IJCAI, 145–154. Rzutowanie na podjęzyk pozwala w wielu przypadkach mapować *wsc* i *snc* na prostsze formuły, wymagające do obliczeń deterministycznego czasu wielomianowego, a w przypadku mapowań na logikę I-go rzędu – dodatkowo znakomitego ograniczenia do logarytmicznej pamięci. Nie znalazłem w rozprawie odniesień do tego kierunku badań.

¹Tworzenie indeksu definiowanych pojęć jest w \LaTeX u bardzo proste. Z pewnością indeks ułatwiłby czytelnikowi rozprawy nawigację po tekście.

Rozdział zawiera także aproksymacyjną semantyką ATL_{ir} , oraz ciekawe dyskusje dotyczące optymalizacji struktur danych i operacji. Podrozdział 3.3. wprowadza nową technikę weryfikacji, wykorzystującą lokalne modele agentów. Proponowane podejście zachowuje kluczowe zależności między agentami dzięki ukierunkowanym kryteriom równoważności obserwacyjnej, co umożliwia efektywną weryfikację zarówno własności bezpieczeństwa, jak i żywotności z określonymi granicami dokładności.

Rozdział jest zakończony ciekawą dyskusją wyzwań i podsumowaniem doświadczeń (po stronie podstaw teoretycznych) zebranych w czasie badań.

Rozdział 4

Czwarty rozdział jest oparty na pracach Jamroga et al. (2022b) oraz Kurpiewski et al. (2019b).

Przedstawiono algorytm syntezy strategii oparty na przeszukiwaniu w głąb, dostosowany do specyfiki problemów epistemicznych niedeterminizmem działania koalicji. Niedeterminizm wymaga uwzględniania zależności między różnymi gałęziami przeszukiwania, co może wymagać zmiany lokalnie poprawnych strategii. Algorytm wykorzystuje mechanizm nawracania, umożliwiającą ich rewizję. Ze względu na ogromną przestrzeń strategii, opracowany algorytm stanowi bazę dla bardziej zaawansowanych technik, które pozwalają efektywnie ją przeszukiwać. Przedyskutowane są też odnośne heurystyki.

Do ciekawszych wyników tego rozdziału zaliczyłbym twierdzenie 4.2.7 o złożoności CO-NP badania optymalności strategii częściowych.

Rozdział 5

Rozdział ten jest oparty na pracach Jamroga and Kurpiewski (2023) oraz Kurpiewski et al. (2019b).

Rozdział dotyczy optymalizacji strategii w systemach wieloagentowych poprzez ich upraszczanie przy zachowaniu gwarancji jakości. Przedstawiono podejścia oparte na dominacji oraz optymalizacji wielokryterialnej, umożliwiające efektywne równoważenie złożoności obliczeniowej i spójności strategicznej. Omówiono także ograniczenia syntezy strategii koalicyjnych oraz zaproponowano metody radzenia sobie z konfliktami celów i problemami epistemicznymi.

Przedstawione rozważania mają głównie charakter praktycznej algorytmiki.

Rozdział 6

Rozdział szósty, dotyczący redukcji przestrzeni stanów, jest oparty na pracach Jamroga et al. (2019a, 2022b).

Problemu eksplozji przestrzeni stanów jest centralny w weryfikacji systemów wieloagentowych. Przedstawiono techniki abstrakcji, grupujące stany i akcje w klasy

równoważności, oraz redukcji częściowego porządku, które eliminują redundantne ścieżki przeszukiwania. Podkreślono kompromis między dokładnością a efektywnością obliczeniową, wskazując, że proponowane metody pozwalają zachować poprawność wniosków przy istotnym zmniejszeniu złożoności.

Podobnie jak w rozdziale drugim nie mogę się oprzeć wrażeniu, że opracowane techniki mieszczą się w dziedzinie zbiorów przybliżonych, czy szerzej obliczeń przybliżonych. Użycie terminu „abstrakcja” (*abstraction*) nie jest przedyskutowane przez Autora. Generalnie – abstrahowanie dotyczy mapowania bazowych systemów na systemy abstrakcyjne, typowo o bardziej abstrakcyjnej sygnaturze (słowniku pojęć), a w konsekwencji – pozwalające na wyrażanie własności na wyższym, bardziej abstrakcyjnym poziomie. Tego aspektu nie zauważyłem w rozprawie, w której abstrakcja jest rozumiana w tradycyjnym matematycznym sensie, gdzie abstrahowanie jest definiowaniem nowych pojęć poprzez klasy abstrakcji (zasada abstrakcji). Tymczasem w sztucznej inteligencji pojęcie abstrahowania jest dużo bardziej ogólne i spodziewałbym się w rozprawie komentarza dotyczącego takiego rozróżnienia.²

Z ciekawych wyników chciałbym zwrócić uwagę na twierdzenie 6.1.1 pokazujące, że dolne i górne aproksymacje zachowują wnioskowanie.

Dalsza część rozdziału dotyczy redukcji opartych na częściowych porządkach.

Rozdział 7

Rozdział siódmy jest oparty na pracach: Jamroga and Kurpiewski (2023); Jamroga et al. (2019a, 2022a,b); Kurpiewski and Marmsoler (2019); Kurpiewski et al. (2019b, 2023).

Proces formalnej weryfikacji wymaga poprawnego sformułowania weryfikowanej własności oraz zdefiniowania modelu obliczeniowego. O ile logiki temporalne dostarczają narzędzi do formalizacji własności, o tyle modele są w dużej mierze tworzone w mniej sformalizowanym procesie, wymagającym eksperckiej wiedzy i odpowiedniej abstrakcji. Kluczowe wyzwania obejmują adekwatność abstrakcji oraz obliczeniową wykonalność. Autor dyskutuje możliwości znalezienia właściwego bilansu między ekspresywnością modelu a możliwościami jego automatycznej jego analizy.

Dyskusja jest przeprowadzona na wybranych studiach przypadku: gry w brydża, zespołów dronów, koordynacji inteligentnych linii produkcyjnych, w których współdziałają maszyny i roboty, czyspołecznie wyjaśnialnej sztucznej inteligencji budowanej z zużyciem metod uczenia maszynowego.

Omówiona jest także technika oparta o modele losowe. Rozdział nie wprowadza

²Nawiasem mówiąc, tematyka abstrahowania poprzez dolne i górne aproksymacje jest omówiona w pracy: Szałas, A. 2026. *Bridge and Bound: A Logic-Based Framework for Abstracting (Extended Report)*. <https://arxiv.org/abs/2510.26654>. Oczywiście Autor rozprawy nie miał do tej pracy dostępu w czasie badań, wspominam tu o niej jedynie by zilustrować swój komentarz.

nowości teoretycznych, czy koncepcyjnych, stanowiąc raczej wprowadzenie do pozostałych rozdziałów.

Rozdział 8

Rozdział ósmy jest oparty o prace Kurpiewski et al. (2019a, 2021).

Rozdział omawia rozwiązania przyjęte w autorskim systemie STV stworzonym do weryfikacji modelowej opartej o techniki i algorytmy przedstawione we wcześniejszych rozdziałach. Modele są specyfikowane jako asynchroniczne systemy wieloagentowe w dobrze zdefiniowanym języku specyfikacji. Rozdział opisuje implementację STV oraz interfejs użytkownika. System stanowi narzędzie do przeprowadzenia badań eksperymentalnych.

Weryfikator STV został zaprogramowany w języku PYTHON. Nie znalazłem w rozprawie dyskusji wyboru języka PYTHON, choć wiadomo, że jest on jest wielokrotnie wolniejszy niż C czy C++ i do nadrobienia prędkości używa się w nim bibliotek napisanych w C/C++.

Niestety nie jest jasne jaka jest rola weryfikatora w przeprowadzonych pracach eksperymentalnych. W rozdziale 9 pojawia się o nim wzmianka tylko w podrozdziale 9.6, gdzie przedstawiono wyniki eksperymentalne problemu głosowania. Wydaje się więc, że weryfikator STV został zbadany eksperymentalnie w pracy w ograniczonym zakresie.

Rozdział 9

Dziewiąty rozdział rozprawy jest oparty o prace: Jamroga and Kurpiewski (2023); Jamroga et al. (2019a, 2022a,b); Kurpiewski and Marmsoler (2019); Kurpiewski et al. (2019b, 2023).

Prace eksperymentalne objęły ewaluację modeli omówionych w studiach przypadku z rozdziału 7. Dokładna weryfikacja modelowa ATL_{ir} została przeprowadzona przy użyciu znanych narzędzi MCMAS, UPPAAL. W celu wykonania weryfikacji przybliżonej wykorzystano jawne reprezentacje modeli z Sekcji 7.6.2 oraz implementację algorytmów wykorzystujących techniki stałopunktowe z Rozdziału 3 w postaci odrębnego narzędzia napisanego w C++.³

Wyniki przeprowadzonych eksperymentów pokazują, że używanie technik aproksymacji i abstrakcji zwykle kończy się sukcesem, podczas gdy weryfikacja dokładna w znacznej większości przypadków kończy się z powodu przekroczenia czasu. Techniki aproksymacyjne i abstrakcji doprowadzają jednak do przekroczenia pamięci przy większych rozmiarach przestrzeni stanów. Warto podkreślić, że w zdecydowanej większości przypadków algorytmy przedstawione w pracy są dużo bardziej wydajne niż algorytmy konkurencyjne (dokładne), co wskazuje na

³Przyznam, że trudno mi było doczytać się które z badań eksperymentalnych zostały przeprowadzone w STV, a które w osobnym narzędziu napisanym w C++.

duży sukces w bardzo wymagającej i konkurencyjnej dziedzinie.

Testy przeprowadzono na procesorze Intel Core i7-6700, 2,60–3,50 GHz, z 32GB RAM, działającym pod kontrolą 64-bitowego systemu Ubuntu Linux. Nie jest to sprzęt bardzo silny, co wykazuje duży potencjał praktycznego wykorzystania opracowanych algorytmów.

3 Ocena rozprawy

Rozprawa zawiera wiele oryginalnych wyników dotyczących algorytmów i metod weryfikacji modelowej. Oryginalne są także badania eksperymentalne. Wkład Autora jest dobrze opisany w rozdziale 1.3: we wszystkich omawianych publikacjach był on znaczący, a w niektórych z nich – wiodący. Uważam, że z nadmiarem stanowi podstawę rozprawy doktorskiej.

Przedstawione wyniki teoretyczne, algorytmiczne, implementacyjne i eksperymentalne są bardzo silne i wskazują na doskonałe opanowanie przez Autora warsztatu badawczego. Wyniki eksperymentalne są dużym osiągnięciem zwłaszcza w świetle teoretycznych wyników złożonościowych przedstawionych w rozprawie (w tym w rozdziale 2).

Do słabszych elementów zaliczyłbym:

- niepełne zbadanie weryfikatora STV: został on użyty tylko w części studiów przypadku, choć wszystkie elementy do bardziej całościowej weryfikacji Autor miał opracowane,
- podejście do abstrakcji zawężone do użycia relacji równoważności: wydaje się, że uwzględnienie innych technik stosowanych w sztucznej inteligencji mogłoby stworzyć przestrzeń do dalszej poprawy wyników eksperymentalnych,
- brak porównania podejścia ze zbiorami przybliżonymi.

Spodziewałbym się też dokładniejszych informacji dotyczących wyboru użycia PYTHON vs C++ oraz bardziej uporządkowanego wskazania narzędzi przy poszczególnych badaniach eksperymentalnych.

Z drobnych uwag – większej staranności oczekiwałbym w formatowaniu tekstu rozprawy.

4 Konkluzja

Mimo wskazanych drobnych niedoskonałości, rozprawa zawiera bardzo ciekawe wyniki teoretyczne i eksperymentalne. Uważam, że spełnia ona wymagania ustawowe i zwyczajowe stawiane przed rozprawami doktorskimi. Wnoszę o jej dopuszczenie do dalszych faz przewodu doktorskiego.

Zakładając pozytywny przebieg obrony wnoszę o wyróżnienie rozprawy ze względu na:

- osadzenie badań w jednym z głównych nurtów badań prowadzonych na świecie oraz połączenie niebanalnej teorii z eksperymentami, przy jednoczesnym uzyskaniu dużo lepszych wyników w ewaluacji rozwiązań,
- duży potencjał opracowanych modeli i algorytmów do dalszego ich rozwoju i zastosowań,
- doskonały dorobek publikacyjny: wyniki wszystkich rozdziałów są poparte publikacjami w źródłach bardzo dobrych lub nieco niższych rangą, ale nadal cieszących się bardzo dobrą reputacją.

**Andrzej
Szałas**  Elektronicznie
podpisany przez Andrzej
Szałas
Data: 2026.03.31
09:26:21 +02'00'