

RECENZJA ROZPRAWY DOKTORSKIEJ DLA RADY NAUKOWEJ
DYSCYPLINY INFORMATYKA TECHNICZNA I TELEKOMUNIKACJA
POLITECHNIKI WARSZAWSKIEJ

Autor recenzji: prof. dr hab. Andrzej Szałas
Wydział Matematyki, Informatyki i Mechaniki
Uniwersytet Warszawski

12 września 2023

AUTOR ROZPRAWY: mgr Teofil Sidoruk

TYTUŁ ROZPRAWY: „State Space Reductions for Multi-agent Systems”

1. WPROWADZENIE

Weryfikacja modelowa (*model checking*) polega na sprawdzaniu czy skończony model systemu informatycznego spełnia zadaną specyfikację. Stanowi ona obecnie jeden z najważniejszych nurtów formalnej weryfikacji systemów począwszy od elementów sprzętowych, poprzez oprogramowanie, po komponenty wysokopoziomowe¹ czy zachowania całych systemów. Opracowane metody są szeroko stosowane w praktyce i wykazały swoją dużą skuteczność. Jedną z barier stosowania zaimplementowanych narzędzi jest jednak złożoność problemu weryfikacji modelowej, wynikająca:

- z samego problemu weryfikacji, trudnego dla bardziej zaawansowanych języków specyfikacji, lokująca się często w klasach złożoności NP, PSPACE, EXPTIME, lub nawet – zależnie od wykorzystywanych języków – wśród problemów nierozstrzygalnych;
- rozmiarów modeli, które nawet specyfikowane w zwarty sposób, wymagają rozwinięcia do dużej liczby stanów, nierzadko prowadząc do ich wykładniczej eksplozji.

Osobnym zagadnieniem jest wybór stosowanych języków specyfikacji własności systemów oraz modeli. Dąży się do uzyskania właściwego balansu pomiędzy siłą wyrazu i zwartością opisową tych języków, a złożonością obliczeniową wyrażonej z ich pomocą problemów weryfikacji modelowej.

W omawianej rozprawie do specyfikacji własności systemów stosuje się warianty/rozszerzenia logik temporalnych, w tym: ATL* (*Alternating-time Temporal Logic*) (a dokładniej – jego ograniczony fragment – SATL*) oraz jej epistemiczny wariant SATLK*. Dodatkowo logiki są uzupełnione o operatory odnoszące się do strategicznych możliwości agentów. Modele są natomiast specyfikowane przez asynchroniczne systemy wieloagentowe AMAS (*Asynchronous Multi-agent Systems*), ich rozszerzenie EAMAS (*Extended AMAS*) oraz drzewa ataku-obrony ADTREES (*Attack-defence Trees*).

W rozprawie Autor postawił tezę:²

„Redukcję częściowych porządków, opracowaną dla liniowej logiki temporalnej LTL, można zaadoptować do weryfikacji strategicznych możliwości agentów w asynchronicznych systemach wieloagentowych. Ponadto wyspecjalizowane techniki mogą dawać silniejsze redukcje, jednakże dla mniejszych klas modeli.”

¹Jak współpracujące ze sobą procesy, podsystemy, roboty, agenci.

²Tłumaczenie autora recenzji (rozprawa jest przedstawiona w jęz. angielskim).

Można stwierdzić, że zagadnienia naukowe i problemy postawione w rozprawie są jasno sformułowane. Ponadto teoretycznym wynikiem rozprawy towarzyszą badania eksperymentalne. Uważam, że zagadnienia te są ciekawe z naukowego punktu widzenia i niewątpliwie warte badania. Z całą pewnością stanowią bardzo dobrą tematykę dla rozprawy doktorskiej.

2. ZAWARTOŚĆ ROZPRAWY

Rozprawa składa się z siedmiu rozdziałów. Pierwszy zawiera wstęp, ostatni – konkluzje. Poniżej skupię się na rozdziałach 2-5, bezpośrednio prezentujących przeprowadzone badania.

Rozdział 2

Rozdział rozpoczyna się od wprowadzenia asynchronicznych systemów wieloagentowych (AMAS) dla n agentów.³ Każdy agent jest pewną maszyną skończenie stanową: ma swój lokalny stan, stan początkowy, zbiór zdarzeń, protokół wskazujący na zdarzenia dostępne w poszczególnych stanach i częściową funkcję przejścia między stanami pod wpływem zdarzeń. Dla interpretacji zachowań agentów wprowadza się dalej przepłotowe systemy interpretowane IIS (*Interleaved Interpreted Systems*), generujące możliwe ścieżki z przepłotami *stan-zdarzenie-stan*, odzwierciedlającymi funkcje przejść agentów. Nad strukturami składającymi się z tych ścieżek definiuje się semantykę logiki ATL*, gdzie obok standardowych operatorów temporalnych można używać modalności postaci $\langle\langle A \rangle\rangle$. Formuła $\langle\langle A \rangle\rangle\gamma$ intuicyjnie oznacza, że zbiór agentów A (koalicja) ma wspólnie strategię gwarantującą własność γ . Dalej Autor rozprawy przedstawia taksonomię strategii oraz pojęcie fair współbieżności, by zakończyć wprowadzeniem epistemicznego wariantu ATLK* logiki ATL* oraz ich ograniczeń SATL*, SATLK*. W tym rozdziale, poza dość prostym stwierdzeniem (Proposition 2.3.2) nie ma innych twierdzeń. Stanowi on jednak istotne wprowadzenie do dalszych rozważań.

Choć w modelowaniu złożonych zjawisk przyjmuje się wiele uproszczeń, wydaje się, że niektóre z tych przyjętych w rozprawie można byłoby dość łatwo pominąć, bez straty jakości uzyskanych wyników. W szczególności pominięty jest ważny dla systemów wieloagentowych aspekt współpracy agentów przy wykonywaniu akcji/powodowaniu zdarzeń. W AMAS oraz IIS do zajścia zdarzenia wystarczy by jeden z agentów miał możliwość jego spowodowania. Efekty wprawdzie uwidaczniają się w zmianie stanów wszystkich agentów mających możliwość spowodowania danego zdarzenia, ale nie można wymusić, by zdarzenie wymagało aktywności więcej niż jednego agenta. Wynika to m.in. ze skupienia się na informacji niedoskonałej (*imperfect information*), rozumianej jako definiowanie zachowań agentów poprzez ich lokalny stan.⁴ Przyjęcie tego założenia jest oczywiście uzasadnione praktyką systemów wieloagentowych, w których zwykle zakłada się dostęp agenta do stanu lokalnego. Niemniej jednak, gdyby modelować zachowania grupowe, w których do zajścia zdarzenia jest niezbędne współdziałanie agentów, należałoby funkcję przejść definiować zależnie od stanów współdziałających agentów. Nadal rozważany byłby stan „prawie lokalny”, ale obejmujący agentów niezbędnych do spo-

³Z formalnego punktu widzenia należałoby napisać, że n jest liczbą naturalną, $n \geq 1$.

⁴Warto podkreślić, że w wielu obszarach sztucznej inteligencji za informację niedoskonałą uważa się informację niepełną, czasem sprzeczną, czy niepewną.

wodowania danego zdarzenia. Odpowiednia modyfikacja definicji w tym kierunku nie wydaje się ingerować negatywnie w dalsze wyniki, w szczególności w metody redukcji przestrzeni stanów.

Niewykluczone, że wiele aspektów współpracy agentów da się wyrazić przy pomocy strategii i użycia operatora $\langle\langle A \rangle\rangle$, rozważanych w rozprawie, ale nie jasne czy i w jakim stopniu jest to możliwe. Zabrakło mi dyskusji możliwości i ograniczeń takiego podejścia.

Rozdział 3

Trzeci rozdział pokazuje jak zaadoptować metodę redukcji częściowych porządków POR (*Partial Order Reduction*) do SATL* z różnymi wariantami semantycznymi oraz operatorem epistemicznym. Rozważa się tu:

- SATL* z informacją niedoskonałą;
- SATL* z informacją niedoskonałą ze współbieżnością niedyskryminującą (*fair concurrency*);
- epistemiczny wariant SATLK*;
- subiektywne możliwości strategiczne.

Wykorzystywany jest algorytm POR, który w swojej istocie polega na generowaniu podmodelu poprzez zawężanie eksploracji typu DFS: Zamiast eksplorować całą przestrzeń stanów – wybiera się heurystycznie odpowiedni podzbiór zdarzeń mogących zajść i pomija pozostałe. Postawą jest jak najlepszy wybór podzbioru zdarzeń by z jednej strony zminimalizować jego rozmiar (co daje możliwie największą redukcję), a z drugiej – by wystarczył dla zachowania rozważanej klasy własności. Znalezienie minimalnego takiego podzbioru jest problemem NP-trudnym i dlatego w pracy rozważa się heurystyki prowadzące do górnej jego aproksymacji.

Wykazane zostały twierdzenia (Tw. 3.4.4, 3.4.5, 3.4.6, 3.4.10, 3.4.12, 3.4.13), które odnoszą się do poprawności redukcji i stanowią bazę teoretyczną uzasadniającą wykorzystanie algorytmu POR.⁵ Są to ciekawe i ważne twierdzenia.

W podrozdziale 3.5 przedstawiono eksperymentalną ewaluację rozważanych metod. Jest ona o tyle istotna, że teoretyczne oszacowanie stopnia uzyskiwanej redukcji nie jest podjęte w doktoracie.⁶ Do eksperymentów wykorzystano otwartoźródłowe narzędzie SPIN. W rozprawie zabrakło mi dyskusji sposobu wykorzystania SPIN. Nie jest jasne, czy wymagało ono adaptacji poprzez zmianę/uzupełnienie kodu, czy udało się zakodować rozważane redukcje w ramach SPIN, bez konieczności jego modyfikacji? Do eksperymentów użyte zostały cztery problemy (opisane w sekcji 3.5.3). Wyniki eksperymentów są bardzo obiecujące – od blisko 100% redukcji stanów i tranzycji po redukcje pomiędzy 40%–70%. Wskazują na duży potencjał zastosowań w praktyce inżynierii tworzenia systemów.

Rozdział 4

Rozdział ten podejmuje zagadnienie redukcji stanów w modelach mniej ogólnych niż częściowe porządki, ale nadal interesujących z punktu widzenia zastosowań. Rozważa się

⁵Przy tym tw. 3.4.4 odnosi się do poprawności pośrednio – jest wykorzystywane przy dowodzeniu innych twierdzeń, może bardziej przejrzyste byłoby nazwać je lematem.

⁶Problem teoretycznego oszacowania stopnia redukcji, nawet przy uproszczonych założeniach, wydaje się tematyką na tyle trudną, że mogłaby stanowić podstawę osobnej rozprawy doktorskiej.

tu znane z literatury drzewa ataku-obrony ADTREES. Dzięki ograniczeniu do drzew można się spodziewać wykorzystania specyficznych własności modeli, nie zachodzących w bardziej ogólnym przypadku dowolnych częściowych porządków. Ograniczenie klasy modeli niejednokrotnie stwarza dodatkowe możliwości. Podobnie okazuje się być w przypadku ADTREES.

W celu osiągnięcia większych redukcji formalizm ADTREES jest tłumaczony na systemy aktualizacji ze strażnikami GUS (*Guarded Update Systems*), które są to skończone stanowymi maszynami, w których tranzycje są uwarunkowane formułami zdaniowymi („strażnikami”) i dodatkowo zawierają aktualizacje wykonywane podczas tranzycji. Następnie rozważa się dwa rodzaje redukcji:

- redukcję opartą na wzorcach (*pattern reduction*), gdzie bardziej złożone wzorce podsystemów zastępuje się prostszymi, ale zachowującymi semantykę;
- redukcję warstwową (*layer-based reduction*), gdzie wykorzystuje się specyficzne własności drzew dla dokonywania redukcji zachowujących osiągalność stanów.

O ile pierwsza z tych redukcji polega na stosowaniu mniejszych podgrafów od tych, które wynikają z bezpośredniej reprezentacji i jest stosunkowo prosta, druga wymaga bardziej zaawansowanej maszynerii teoretycznej. Poświęcony jest jej podrozdział 4.6, którego najważniejszy wynik jest formułowany jako stwierdzenie 4.6.11 o zachowywaniu osiągalności. W uzasadnieniu używa się lematu 4.6.7. Wydaje się, że dowód 4.6.11 powinien być nieco bardziej rozbudowany – w obecnej postaci jest nadmiernie skrótowy. Nawiasem mówiąc nie widać wykorzystania lematu 4.6.8. Warto byłoby wskazać gdzie są wykorzystane własności ukierunkowania korzeniowego (*root-directedness*) i odseparowania aktualizacyjnego (*update-separateness*) ADTREES.

W podrozdziale 4.7 przedstawiono weryfikację eksperymentalną:

- w sekcji 4.7.2 na podstawie wcześniej dyskusowanego scenariusza *treasure hunters* oraz trzech scenariuszy zaczerpniętych z literatury; wyniki są przedstawione w tabeli 4.3;
- w podrozdziale 4.7.3 na podstawie generowanych ADTREES; wyniki są przedstawione w tabeli 4.4.

Wyniki eksperymentów są ponownie bardzo obiecujące i wskazują na duży potencjał opracowanej metody (ADTREES są wykorzystywane w literaturze przedmiotu przez wielu autorów). Do pokazania działania algorytmów wykorzystano scenariusze z rozdziału 4 oraz kilka innych, mniejszych scenariuszy.

Rozdział 5

W rozdziale 5 przeprowadza się redukcję przestrzeni stanów poprzez zmniejszanie liczby agentów, zorientowaną na scenariusze bezpieczeństwa, a więc dziedzinę badań, której rolę trudno obecnie przecenić. W tym celu najpierw rozszerza się model AMAS tak, by reprezentował ADTREES. Uzyskany model EAMAS zawiera dodatkowo takie elementy, jak zdaniowe formuły stanowiące „strażników” tranzycji, oraz funkcję aktualizacji. W sekcji 5.2.2 pokazana jest translacja ADTREES na EAMAS, tym razem z pominięciem GUS. Podsumowaniem translacji są twierdzenia 5.2.3 i 5.2.4, wskazujące na jej pełność i poprawność.

Przejście na formalizm EAMAS pozwala na używanie agentów w scenariuszach związanych z bezpieczeństwem. Choć to naturalny krok – dotychczas nie był badany przez

innych autorów w kontekście redukcji stanów i tranzycji. Zaproponowano algorytm minimalizacji rozmiarów grup agentów tak, aby przy odpowiednim przydziale zadań mogły one skutecznie wykonać atak. Generowaniu optymalnego planu jest poświęcony podrozdział 5.4 i podane w nim algorytmy. Zbadano złożoność i poprawność algorytmów (stwierdzenia 5.4.3 i 5.4.4). Wydaje się, że stwierdzenie (Proposition) 5.4.4 zasługuje na status twierdzenia.

W podrozdziale 5.5 przedstawiono wyniki eksperymentów, pokazujących działanie algorytmów i testujących ich poprawność.

3. CZY W ROZPRAWIE PRZEPROWADZONO W SPOSÓB WŁAŚCIWY ANALIZĘ ŹRÓDEŁ (W TYM LITERATURY ŚWIATOWEJ, STANU WIEDZY I ZASTOSOWAŃ W PRZEMYŚLE) ŚWIADCZĄCY O DOSTATECZNEJ WIEDZY AUTORA? CZY WNIOSKI Z PRZEGLĄDU ŹRÓDEŁ SFORMUŁOWANO W SPOSÓB JASNY I PRZEKONYWUJĄCY?

Każdy rozdział, poza pierwszym (wstępnym) i ostatnim (z konkluzjami), zawiera sekcję „Related work”. Każda z tych sekcji w sposób właściwy sytuuje uzyskane wyniki w kontekście literatury przedmiotu. Ponadto w tekście znajduje się szereg cytowań do prac innych autorów oraz wyników własnych, opublikowanych w artykułach współautorskich (6 wchodzących w skład rozprawy oraz 3 inne,⁷ mające jednak dość bliski związek z tematyką rozprawy). W rozprawie zostało zacytowanych 100 źródeł. Wiedza Autora w szeroko rozumianej tematyce rozprawy jest więc obszerna.

Ponadto każdy z rozdziałów z wynikami (rozdz. 2–5) zaczyna się od wskazania prac własnych (współautorskich), będących punktem wyjścia do dalszych części rozdziału oraz roli Autora rozprawy w raportowanych tam badaniach.

4. CZY AUTOR ROZWIĄZAŁ POSTAWIONE ZAGADNIENIA, CZY UŻYŁ WŁAŚCIWEJ DO TEGO METODY I CZY PRZYJĘTE ZAŁOŻENIA SĄ UZASADNIONE?

Autor rozwiązał postawione zagadnienie zarówno od strony teoretycznej, jak i eksperymentalnej, poprzez:

- opracowanie rozszerzenia istniejących metod na nowe formalizmy, wymagające przeprowadzenia badań – wynikiem jest szereg nowych definicji formalizmów i odnoszących się do nich twierdzeń;
- przeprowadzenie eksperymentów wskazujących na skuteczność opracowanych metod.

Użyte metody są właściwe – opracowane formalizmy są rozszerzeniami formalizmów szeroko stosowanych przez innych badaczy. Metodologia badań teoretycznych jest oparta o bardzo dobrze ugruntowane podstawy logiczne. Eksperymenty też zostały opracowane w sposób właściwy.

Przeprowadzone badania w pełni uzasadniają zadanie badawcze przedstawione w rozprawie (zacytowane w niniejszej recenzji na końcu pierwszej strony).

⁷W tym – co warto podkreślić – dwie na AAMAS’2021, AAMAS’2023. Konferencja AAMAS (ranga A*) jest najważniejszą światową konferencją w dziedzinie systemów wieloagentowych.

5. NA CZYM POLEGA ORYGINALNOŚĆ ROZPRAWY, CO STANOWI SAMODZIELNY I ORYGINALNY DORÓBEK AUTORA, JAKA JEST POZYCJA ROZPRAWY W STOSUNKU DO STANU WIEDZY CZY POZIOMU TECHNIKI REPREZENTOWANYCH PRZEZ LITERATURĘ ŚWIATOWĄ?

Rozprawa przedstawia wiele oryginalnych wyników, opublikowanych w 6 artykułach. Wprawdzie są one współautorskie, jednak Autor rozprawy miał w nich istotny udział poprzez nowe (rozszerzone w stosunku do znanych z literatury) definicje, twierdzenia (w szczególności przeprowadzenie pełnych ich dowodów) oraz przeprowadzenie eksperymentów.⁸

Publikacja wyników w bardzo dobrych źródłach (zwłaszcza *J. Artif. Intell. Research* i konferencja *KR'2021*) pokazują uznanie oryginalności i wagi wyników w środowisku międzynarodowym.

6. CZY AUTOR WYKAZAŁ UMIEJĘTNOŚĆ POPRAWNEGO I PRZEKONYWUJĄCEGO PRZEDSTAWIENIA UZYSKANYCH PRZEZ SIEBIE WYNIKÓW (ZWIĘZŁOŚĆ, JASNOŚĆ, POPRAWNOŚĆ REDAKCYJNA PRACY)?

Autor wykazał opanowanie warsztatu badawczego, w tym umiejętność poprawnego i przekonującego przedstawiania swoich wyników zarówno w warstwie teoretycznej, jak i eksperymentalnej. Poza samą rozprawą świadczą o tym także artykuły stanowiące jej podstawę, opublikowane w bardzo dobrych źródłach.

7. JAKIE SĄ SŁABE STRONY ROZPRAWY I JEJ GŁÓWNE WADY?

Choć rozprawa jest bardzo interesująca i zawiera szereg ważnych wyników, zabrakło mi szerszych dyskusji dotyczących:

- dokonanych wyborów związanych z modelowaniem (np. dotyczących współpracy grupowej, czy wyboru modelu przepłotowego);
- adaptacji niezbędnych do wykorzystania narzędzia SPIN;
- dyskusji skalowalności rozwiązań (np. dla jakich rozmiarów problemów nie udaje się uzyskać redukcji ze względu na zbyt dużą przestrzeń stanów do eksploracji nawet przy stosowaniu dobrych heurystyk).⁹

Z usterek redakcyjnych zwraca uwagę brak źródeł przy rysunkach zaczerpniętych z innych prac, w tym takich, których Autor rozprawy nie jest współautorem (np. rys. 2.1, 2.2, 3.4 pochodzą z artykułu [25]).

8. JAKA JEST PRZYDATNOŚĆ ROZPRAWY DLA NAUK TECHNICZNYCH?

Rozprawa dotyczy ważnego zagadnienia weryfikacji modelowej, stosowanej w inżynierii systemów informatycznych. Opracowane algorytmy poprawiają skalowalność znanych rozwiązań poprzez redukcję rozmiarów modeli, co ma poważne znaczenie dla zakresu zastosowań narzędzi weryfikacji modelowej. Ponieważ odnośne techniki są używane w wielu dziedzinach inżynierskich, przydatność rozprawy dla nauk technicznych uważam za dużą.

⁸Autor rozprawy opisuje swoją rolę w powstaniu artykułów na początku każdego z rozdziałów 2–5.

⁹Jedynie w tabelach 4.3, 4.4 dotyczących redukcji ADTREES pojawia się brak wyników ze względu na przekroczenie czasu obliczeń.

9. PYTANIA DO AUTORA ROZPRAWY

W czasie czytania rozprawy nasunęło mi się kilka pytań, na które chciałbym usłyszeć odpowiedzi podczas obrony:

1. Czy i w jakim zakresie można wyrazić współpracę agentów poprzez strategie lub inne elementy rozważanych formalizmów (np. wymóg by do zajścia zdarzenia było potrzeba więcej niż jednego agenta)?
2. Jaki wpływ na wyniki miałyby dopuszczenie w modelu AMAS niedeterminizmu (zamiast funkcji lokalnej tranzycji T_i w def. 2.1, rozważałoby się relację tranzycji)?
3. Czy Autor rozprawy rozważał problem znalezienia klas własności, dla których problem wyznaczenia minimalnego zbioru zdarzeń w algorytmie POR staje się wielomianowy?
4. Czy narzędzie SPIN wymagało adaptacji/rozszerzeń do przeprowadzenia eksperymentów? Jeśli tak, to jakich?
5. Jakie są ograniczenia stosowalności opracowanych metod? Dla jakich rozmiarów modeli nie udaje się przeprowadzić redukcji? Np. dla jak dużych wartości parametrów n rozważanych w sekcji 3.5.4 następuje przekroczenie czasu?

10. DO KTÓREJ Z NASTĘPUJĄCYCH KATEGORII RECENZENT ZALICZA ROZPRAWĘ:
- (A) NIE SPEŁNIAJĄCA WYMAGAŃ STAWIANYCH ROZPRAWOM DOKTORSKIM PRZEZ OBOWIĄZUJĄCE PRZEPISY
 - (B) WYMAGAJĄCA WPROWADZENIA POPRAWEK I PONOWNEGO RECENZOWANIA
 - (C) SPEŁNIAJĄCA WYMAGANIA
 - (D) SPEŁNIAJĄCA WYMAGANIA Z WYRAŹNYM NADMIAREM
 - (E) WYBITNIE DOBRA, ZASŁUGUJĄCA NA WYRÓŻNIENIE.

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.

Ze względu na:

- połączenie niebanalnej teorii z eksperymentami, wykazujące bardzo solidny warsztat badawczy Autora;
- osadzenie badań w jednym z głównych nurtów badań prowadzonych na świecie nad weryfikacją modelową;
- duży potencjał opracowanych modeli i algorytmów do dalszego ich rozwoju, badań i zastosowań;
- fakt, że wyniki wszystkich rozdziałów są wsparte publikacjami w źródłach bardzo dobrych (lub nieco niższych rangą, ale nadal cieszących się bardzo dobrą reputacją),

zaliczam rozprawę do kategorii (E) i wnoszę o jej wyróżnienie.



ANDRZEJ SZALAS