Politechnika Wrocławska

Prof. dr hab. inż. Halina Kwaśnicka
Katedra Sztucznej Inteligencji
Wydział Informatyki i Telekomunikacji
Politechnika Wrocławska
Wyb. Wyspiańskiego 27, 50 370 Wrocław
E-mail: halina.kwasnicka@pwr.edu.pl

18 października 2024r.

Recenzja rozprawy doktorskiej

Tytuł rozprawy: AFFINE CONTINUOUS CELLULAR AUTOMATA AND THEIR ROLE IN SOLVING DENSITY CLASSIFICATION PROBLEMS

Autor rozprawy: mgr Marcin Dembowski

Promotor: prof. Bernard DE BAETS
Promotor pomocniczy: dr hab. Barbara Wolnik

Recenzja wykonana jest na zlecenie Rady Naukowej Instytutu Badań Systemowych PAN w dyscyplinie *Informatyka techniczna i telekomunikacja*

1. Obszar problemowy pracy

Obszar badań prezentowanych w recenzowanej pracy to automaty komórkowe (CA, ang. Cellular Automata). Autor skupia się na problemie klasyfikacji gęstości (DCP, ang. Density Classification Problem). Jest to popularne zadanie w dziedzinie automatów komórkowych, polega na stworzeniu reguły (najczęściej w postaci automatu komórkowego), która potrafi skategoryzować całą sieć jako mającą przewagę jedynek bądź zer, na podstawie tylko wstępnych danych binarnych. Sieć, o której tu mowa, to np. tablica komórek, jedno- bądź wielowymiarowa. Automat komórkowy ma tylko lokalną wiedzę, o danej komórce i jej sąsiedztwie, dlatego problem DCP jest trudny. Automat komórkowy, w oparciu o skonstruowaną regułę powinien podejmować decyzję na podstawie lokalnej informacji tak, aby osiągnąć globalny cel. DCP może być interesujący z praktycznego punktu widzenia. Jednocześnie należy stwierdzić, że nie istnieje idealna reguła jednowymiarowa dla automatu komórkowego, która będzie rozwiązywać ten problem niezawodnie, w każdym przypadku. Z tego powodu rozważa się rozluźnione, tzw. relaksacyjne problemy DCP (ang. Relaxed DCP), w których wymagania problemu klasyfikacji gęstości są ‘poluzowane’. W tych problemach dopuszcza się pewien margines błędu. Przykładowo, rozwiązanie może być akceptowalne, jeśli nie wszystkie, ale większość komórek przyjmie poprawny stan. Omawiany problem jest interesujący i może być użyteczny, np. do analizy samoorganizujących się systemów, czy do

lepszego zrozumienia jak złożone systemy mogą rozwiązywać problemy na podstawie tylko lokalnych interakcji.

Autor rozprawy omawia różne automaty komórkowe, elementarne jednowymiarowe oraz dwuwymiarowe automaty komórkowe o binarnych wartościach komórek, ale też ciągłe automaty komórkowe (ang. Continuous Cellular Automata, CCA). CCA są nazywane też rozmytymi automatami komórkowymi (ang. Fuzzy CA), jeśli zbiór stanów CCA jest ograniczony do przedziału jednostkowego [0, 1].

Rozważana w rozprawie grupa automatów komórkowych to afiniczne, ciągłe automaty komórkowe (ang. Affine Continuous Cellular Automata, ACCA). ACCA łączą cechy ciągłych automatów komórkowych z transformacjami afinicznymi, które są wykorzystywane do modyfikowania stanów komórek w sposób ciągły. Zasady zmiany stanów uwzględniają sąsiedztwo komórek, ale zmiany są wyrażane w postaci operacji matematycznych. ACCA mogą być użyteczne do modelowania procesów fizycznych, jak przepływ cieczy czy propagacja ciepła.

Podsumowując tę część rozprawy stwierdzam, że podjęta w pracy tematyka badań jest ważna i rozwijana w światowym środowisku naukowym.

Elementem, którego mi brakuje w recenzowanej pracy, to sformułowanie we wstępie celu i zakresu badań. Oceniając rozprawę naukową zwykle biorę pod uwagę postawiony cel i określony zakres pracy oraz jakość prezentowanych badań. Tu mogę ocenić, że przedstawione badania i ich wyniki są ważnym obszarem, cieszącym się zainteresowaniem badaczy na świecie.

2. Kompozycja i zawartość pracy

Praca stanowi spójny i logiczny opis badań Kandydata. Pierwszy rozdział to historyczne spojrzenie na obszar automatów komórkowych, sięgające do Stanisława Ulama. Prawdę mówiąc, czytając słowo wstępne do rozdziału sądziłam, że autor rozprawy skupi się na zastosowaniach CA do modelowania epidemii. Nic takiego nie ma miejsca. Z uwagi na fakt, że nigdzie nie znalazłam zdefiniowanych celu i zakresu rozprawy doktorskiej, postawionych problemów badawczych, będących jądrem doktoratu, nie wiedziałam czego spodziewać się w dalszej części rozprawy. Jak rozumiem, Autor w rozdziale pierwszym chciał pokazać niezorientowanemu czytelnikowi, że automaty komórkowe to nie tylko teoretyczna ciekawostka, ale że znajdują praktyczne zastosowanie.

Rozdział drugi jest istotny, bowiem definiuje pojęcia i stosowaną w pracy notację. Praca dotyczy skończonych CA jedno bądź dwuwymiarowych, na siatce prostokątnej. Zbiór wartości, jakie mogą przyjmować komórki jest ograniczony, są to binarne CA bądź CCA, w których komórka może przyjmować wartości rzeczywiste z zakresu [0, 1].

Rozdział trzeci wprowadza do Afinicznych Ciągłych Automatów Komórkowych (ang. Affine Continuous Cellular Automata, ACCA) rozszerzając definicje i notacje na tę właśnie grupę. Autor ogranicza rozważania do ACCA o wartościach komórek z zakresu [0, 1].

Rozdział czwarty prezentuje kluczowy problem rozwiązywany w tej pracy, czyli DCP. Autor sformułował klasyczną wersję DCP oraz jego relaksacyjną wersję (ang. Relaxed DCP). Po wprowadzeniu klasycznej wersji DCP dla CCA, Doktorant wyjaśnia, że nie ma ACCA, który byłby w stanie rozwiązać tę wersję DCP. Następnie, Doktorant sformułował rozluźnioną wersję DCP (Relaxed DCP). Rozdział pokazuje, że klasyczną definicję problemu klasyfikacji gęstości można przeformułować przyjmując relaksację, co pozwala na włączenie do badań symulacji komputerowej.

Rozdział piąty poświęcony jest analizie dynamiki podzbioru afinicznych, ciągłych automatów komórkowych (ACCA) – jednowymiarowych, zewnętrznie totalistycznych i legalnych (ang. outer-totalistic and legal) ACCA o promieniu 1. Autor skupia się na symulacji ACCA zaczynających się od konfiguracji binarnych, co oznacza skończony zbiór początkowych konfiguracji.

Rozdział szósty pokazuje, że można uzyskać rodzinę ACCA będących poprawnymi rozwiązaniami zrelaksowanej wersji DCP dla dowolnej początkowej konfiguracji, niezależnie od liczby komórek. Autor omówił badania symulacyjne zachowań jednowymiarowych ACCA, z jednostkowym promieniem sąsiedztwa, w rozwiązywaniu zrelaksowanego DCP.

Rozdział siódmy dedykowany jest rozwiązywaniu jednowymiarowego DCP o stałej długości (ustalonej liczbie komórek), stosując ACCA.

Kolejne dwa rozdziały rozszerzają badania na zrelaksowane dwuwymiarowe DCP. W rozdziale ósmym Autor szuka odpowiedzi na pytanie, które dwuwymiarowe ACCA, zachowujące gęstość, działają najlepiej. W dwóch wymiarach nie istnieje reguła, która działałaby w jednym wymiarze, jak np. reguła ruchu drogowego. Dochodzi do tego duża złożoność obliczeniowa, co utrudnia bądź uniemożliwia przeprowadzenie wszystkich symulacji, jakie są możliwe dla jednowymiarowego przypadku. Tutaj Autor koncentruje badania na szukaniu reguł, które są lepsze niż reguła uśredniania. Z uwagi na złożoność obliczeniową, warte podkreślenia jest spostrzeżenie, że aby uzyskać zadowalające rozwiązanie, wystarczy zbadać jednowymiarową podprzestrzeń reguł obrotowo-symetrycznych.

Rozdział dziewiąty dotyczy rozwiązywania dwuwymiarowego DCP o stałej długości. Tu Autor postanowił wykorzystać podejście ewolucyjne. Przy tak dużych przestrzeniach wydaje się sensowne zastosowanie metaheurystyk. Natomiast zupełnie nie rozumiem dlaczego Autor zamieszcza podrozdział *9.2 History*, pisze w nim o historii obliczeń ewolucyjnych, programowaniu ewolucyjnym (EP), strategii ewolucyjnej (ES), po czym – stosuje ewolucję różnicową. Celowe byłoby zacząć od zdefiniowania DCP jako zadania optymalizacji, uzasadnić stosowanie metaheurystyki, omówić w jaki sposób zastosowano ewolucję różnicową. Zwłaszcza, że np. z rysunkiem Figure 9.1 można dyskutować, bo po inicjalizacji

JK

(wygenerowaniu populacji początkowej) następuje selekcja. Pytanie – na jakiej podstawie, skoro ta populacja nie była oceniona? Dyskusyjne jest też stwierdzenie, że era obliczeń ewolucyjnych zaczęła się od pierwszego programowalnego komputera cyfrowego, ENIACa. Jakie paradygmaty ewolucji biologicznej dojrzał Autor rozprawy w pierwszym komputerze programowalnym pozostaje tajemnicą. Rozdział zawiera omówienie szerokiego wachlarza eksperymentów z wykorzystaniem ewolucji różnicowej, dla różnych rozmiarów dwuwymiarowej siatki. Rozdział dziesiąty stanowi podsumowanie pracy oraz sugeruje kierunki dalszych badań.

Podsumowując tę część recenzji stwierdzam, że struktura recenzowanej pracy jest spójna i logiczna. Autor wprowadza w problematykę automatów komórkowych, definiuje klasyczny problem klasyfikacji gęstości (DCP), a następnie jego wersję relaksacyjną. W dalszej części Kandydat skupia się na stosowaniu afinicznych ciągłych automatów komórkowych do problemów jednowymiarowych, po czym rozszerza badania na problemy dwuwymiarowe.

Wszystkie pojęcia są jasno zdefiniowane, stosowana jest stała notacja, co ułatwia czytanie pracy. W rozprawie podano wiele formalnych definicji, wykazano wiele twierdzeń. Eksperymenty symulacyjne weryfikują wyniki analityczne oraz ukazują dynamikę badanych CA w rozwiązywaniu DCP.

Wydawałoby się, że praca stanowi wzór dysertacji doktorskich. Jednak tak nie uważam. Jak wskazuje punkt 4b. dokumentu *Szczegółowy tryb postępowania w sprawie nadania stopnia doktora przez Radę Naukową IBS PAN*, przedmiotem rozprawy doktorskiej jest „oryginalne rozwiązanie problemu naukowego, oryginalne rozwiązanie w zakresie zastosowania wyników własnych badań naukowych w sferze gospodarczej lub społecznej”. Jest to oczywiste wymaganie, dlatego w każdej rozprawie powinien być zdefiniowany cel i zakres pracy – jaki problem badawczy jest przedmiotem rozprawy. Osobiście jestem przeciwniczką formułowania hipotezy badawczej, przynajmniej w doktoratach bronionych w naukach technicznych. Ale oczekuję jasno sformułowanego celu / problemu badawczego. Czytając recenzowaną pracę, długo nie wiedziałam jaki problem, w jakim zakresie pan Dembowski rozwiązuje. Trochę czekałam na modelowanie epidemii, trochę na praktyczny problem, do którego rozwiązania będą zastosowane różne automaty komórkowe. Co prawda, tytuł rozprawy mówi że chodzi o rolę ACCA w rozwiązywaniu DCP, ale dopiero przy czytaniu drugiej części rozprawy wydawało mi się, że wiem czego się spodziewać dalej. Chociaż i tu zostałam zaskoczona obliczeniami ewolucyjnymi. Podkreślam, że zastosowanie obliczeń ewolucyjnych do tego problemu uważam za w pełni zasadne.

3. Oryginalność i waga osiągnięć

Recenzowana rozprawa dotyczy wciąż aktualnego problemu. Baza WoS, na pytanie „continuous cellular automata” zwraca ponad 1200 pozycji (po dodaniu ‘affine’ jest tylko 10 pozycji). W odpowiedzi na zapytanie „density classification problem” jest to ponad 4700 pozycji (dane z 17.10.2024). Pan Marcin Dembowski wykazuje się dużą sprawnością

w prowadzeniu badań i bardzo dużą wiedzą w tematyce prowadzonych badań. Praca skupia się na szczególnym typie automatów komórkowych, na ciągłych, afinicznych CA w kontekście rozwiązania problemu klasyfikacji gęstości. Przeprowadzone eksperymenty pozwoliły ujawnić bądź zweryfikować pewne właściwości ACCA.

Badania symulacyjne 'legal outer-totalistic' ACCA, które są najprostszym uogólnieniem ECA, wykazały, że przestrzeń reguł można podzielić na szereg klas o odmiennym zachowaniu. Obserwowana dynamika tych automatów znalazła potwierdzenie w postaci dowodów analitycznych.

W pracy pokazano, że istnieje klasa ACCA zachowująca gęstość, rozwiązująca relaksacyjne DCP, czego nie można powiedzieć o CA. Eksperymenty pokazały, że czas konwergencji takich klasyfikatorów zależy od liczby komórek w konfiguracji początkowej oraz od wybranego progu, przy czym dla progu relacja ta nie jest liniowa. Analiza teoretyczna DCP o stałej długości, w kontekście ACCA pozwoliła na ograniczenie symulacji (tu, do rozmiaru $N = 23$). Zbiory rozwiązań zawierają zróżnicowane reguły pod względem czasu konwergencji.

Dla przypadków dwuwymiarowych problem stanowi złożoność obliczeniowa. Okazało się, że można otrzymać zadowalające rozwiązanie przy badaniu jednowymiarowej podprzestrzeni reguł obrotowo-symetrycznych zamiast całej przestrzeni reguł zachowania gęstości opisanych przez 8-wymiarowy wielokąt.

Stosując metaheurystykę, przy jej dobrym dostrojeniu, można znaleźć reguły ACCA rozwiązujące dwuwymiarowy klasyczny DCP, dla danego rozmiaru siatki. Warto ograniczyć się do reguł obrotowo-symetrycznych, zmniejsza to złożoność problemu i dla badanego rozmiaru daje dobre wyniki. Okazało się, że przy stosowaniu ewolucji różnicowej, wykorzystanie reguł ACCA rozwiązujących problem dla mniejszych siatek zmniejszyło liczbę iteracji potrzebną do znalezienia dobrego rozwiązania dla siatki większej.

Analityczne i symulacyjne badania nadal pozostawiają wiele wyzwań. Takim wyzwaniem jest proces weryfikacji reguły ACCA względem dowolnej konfiguracji $n \times n$. Liczba konfiguracji, które należy zweryfikować szybko wzrasta wraz z rozmiarem siatki.

Trudno jest mi ocenić rangę udowodnionych twierdzeń. Jest ich dużo, niektóre z nich Autor zamieszcza, jak sam pisze, dla wygody czytelnika, niektóre są oryginalne. Ale nawet te oryginalne, powstały we współpracy z zespołem. Trudno mi tu podkreślić które twierdzenia czy lematy są największym, teoretycznym osiągnięciem Kandydata. Doktorant, w odróżnieniu od wielu znanych mi rozpraw doktorskich, nie wskazuje, co, Jego zdaniem, jest najważniejszym, oryginalnym, autorskim osiągnięciem. Patrząc na poszczególne rozdziały prezentujące efekty badań, to należy stwierdzić, że wszystkie wyniki są ważne. Przy braku sformułowanego celu i zakresu badań, też nie mogę odnieść się do stopnia, w jakim postawiony problem został rozwiązany czy osiągnięty jest cel rozprawy. Mogę ocenić zakres badań oraz sposób ich poprowadzenia i omówienia wyników.

JK

Podsumowując tę część recenzji stwierdzam, że wszystkie eksperymenty, które, jak wynika z treści rozprawy, wykonywał Autor samodzielnie, są przeprowadzone w prawidłowy sposób. Ich zaplanowanie, przeprowadzenie, prezentacja i analiza wyników oraz formułowane konkluzje nie budzą u mnie żadnych wątpliwości. Widoczna jest duża biegłość Doktoranta w posługiwaniu się językiem formalnym. Wiele dowodów twierdzeń, wykonanych we współpracy z bardziej doświadczoną osobą, również świadczy na korzyść Doktoranta. **Nie mam wątpliwości, że pan Marcin Dembowski zasługuje na uznanie Jego badań i dopuszczenie do dalszego etapu w procesie doktoryzowania**.

4. Uwagi i problemy do dyskusji

Nie mam konkretnych uwag merytorycznych do pracy. Można dyskutować, dlaczego akurat ewolucję różnicową wybrał doktorant do rozwiązania problemu optymalizacji przy złożonych problemach dwuwymiarowych. Nie jest to zarzut, bo na pytanie, jaką metaheurystykę wybrać, odpowiedziałabym: tę, którą najlepiej znasz i czujesz.

Wartością dodaną byłoby rozwiązanie chociaż jednego, rzeczywistego problemu. Czy DCP może być przydatne w rozpoznawaniu wzorców w systemach wizyjnych, np. w medycynie do analizy (czarno-białych) obrazów rentgenowskich lub MRI?

Można wskazać wiele obszarów, gdzie analizowana problematyka może znaleźć zastosowanie. Czy doktorant zna konkretne zastosowanie DCP, wykorzystane w praktyce?

5. Konkluzja

Po zapoznaniu się z rozprawą doktorską, ocenie zaprezentowanych w niej badań analitycznych i symulacyjnych, stwierdzam, iż przedłożona mi do recenzji rozprawa, której autorem jest mgr Marcin Dembowski, **spełnia ustawowe wymagania stawiane rozprawom doktorskim i wnioskuję o dopuszczenie pana Marcina Dembowskiego do publicznej obrony.**



Halina Kwaśnicka