

Warszawa, 24 sierpnia 2018

dr hab. inż. Jarosław Arabas, prof. nzw. PW  
Politechnika Warszawska  
Wydział Elektroniki i Technik Informacyjnych  
Instytut Informatyki  
jarabas@elka.pw.edu.pl

# RECENZJA

rozprawy doktorskiej mgr. inż. Pawła Liskowskiego  
pt. „Heuristic algorithms for discovery of search objectives  
in test-based problems”

## 1. Problem badawczy i jego znaczenie

Dominujące tendencje rozwojowe w optymalizacji numerycznej sprowadzają się do poszukiwania metod optymalizacji, które będą w stanie coraz mniejszym kosztem i z coraz większą skutecznością znajdować przybliżenia optimum funkcji celu. Jego znalezienie uznawane jest za sukces.

Jednak z punktu widzenia praktycznego, sukcesem jest rozwiązanie problemu, a nie jego abstrakcji sformułowanej jako funkcja celu. Autor próbuje zmierzyć się z tym zagadnieniem poprzez próbę uelastyczenia funkcji celu. Zakres rozprawy obejmuje zagadnieniu realizacji maszynowego uczenia poprzez optymalizację wskaźnika jakości. Autor koncentruje się na maszynowej indukcji programów za pomocą techniki programowania genetycznego.

W mojej ocenie postawienie problemu jest bardzo cenne, a jego uszczegółowienie jest zręcznym pomysłem wykorzystującym specyfikę zarówno zadania jak i metody optymalizacji. Wprawdzie wyniki nie mogą być wprost uogólnione na wszystkie możliwe zadania optymalizacji, jednak rozważana klasa zadań bazujących na testach jest dostatecznie szeroka aby potraktować ją jako interesujący obiekt badań. Z punktu widzenia algorytmów ewolucyjnych, lub szerzej, populacyjnych metod optymalizacji, uważam że koncept oceny populacji jako całości jest bardzo nośną i zręczną formalizacją istniejącego w praktyce stanu rzeczy, gdy prawdopodobieństwo selekcji konkretnego osobnika zależy zawsze od wartości funkcji celu wszystkich pozostałych osobników populacji.

Ocenianą rozprawę zaliczam do badań podstawowych w naukach technicznych, w dyscyplinie informatyka. Autor potwierdza swoje rozważania poprzez eksperymenty z użyciem akademickich zadań testowych. Uważam, że rozprawa jest utrzymana na bardzo wysokim poziomie merytorycznym.





## 2. Wkład rozprawy w rozwój nauki

W rozważanej przez Autora klasie zagadnień, wskaźnik jakości używany do uczenia jest typowo funkcją, przyporządkowującą każdemu programowi wartość liczbową, którą można utożsamiać z błędem. Wartość ta jest obliczana jako wyrażenie, najczęściej suma, agregujące wartości błędów popełnianych przez model dla poszczególnych scenariuszy użycia programu, zwanych przez Autora testami. Autor zauważa, że agregacja wartości ocen dla poszczególnych testów powoduje utratę informacji o zachowaniu programu dla *konkretnych* testów, co utrudnia porównywanie dwóch różnych programów. W rozprawie ten efekt jest nazwany *evaluation bottleneck*. Jako remedium Autor proponuje wprowadzenie oceny wielokryterialnej i wykorzystanie jej w procesie poszukiwań rozwiązania przez algorytm ewolucyjny. Jednocześnie wskazane zostaje, że bezpośrednie bazowanie na wartościach testów wprowadzi tak wielką liczbę kryteriów, że większość lub nawet wszystkie programy staną się niezdominowane i proces uczenia nie będzie postępować.

Kluczowym obiektem, który służy do sformułowania nowych koncepcji oceny rozwiązań, jest macierz **G** opisująca wartość uzyskaną dla każdego testu przez każdy z programów, podlegających ocenie. Autor przyjmuje przy tym, że grupą ocenianych jednocześnie programów jest populacja, przetwarzana przez programowanie genetyczne.

W rozprawie rozważone są zasadniczo dwie koncepcje przetworzenia macierzy **G**: wybór niewielkiego podzbioru testów oraz generacja kryteriów oceny na podstawie przybliżonej faktoryzacji macierzy **G** na czynniki małego rzędu (dwa lub trzy). Po wykonaniu tego zabiegu, selekcja przebiegać może zarówno na podstawie oceny wielokryterialnej jak i na podstawie zastępczej funkcji celu, sformułowanej na podstawie zredukowanej liczby kryteriów oceny. Główna idea rozprawy została przez Autora przedstawiona w rozdziale 8, zaś następujące po nim rozdziały 9–12 są poświęcone różnym wariantom jej realizacji:

- rozdział 9: wybór testów na podstawie analizy skupień (metoda DOC),
- rozdział 10: faktoryzacja macierzy i wyłonienie „sztucznych testów” jako wyników faktoryzacji (metoda DOF),
- rozdział 12: uogólnienie DOF na przypadek, w którym nie wszystkie testy zostały wykonane dla ocenianych osobników; dzięki faktoryzacji możliwa jest imputacja wartości nieoznaczonych i użycie ich do sformułowania zastępczej funkcji celu (metoda SFIMX),
- rozdział 11: uogólnienie metod DOC i DOF na przypadek regresji funkcji ciągłej,

W rozdziałach 9,10,12 Autor zweryfikował eksperymentalnie rozważane modyfikacje z użyciem zadań testowych o charakterze binarnym. Były to zadania dwóch rodzajów:

- oparte na grze *reversi* (inaczej *Othello*) oraz iterowanym dylemacie więźnia, w których optymalizowany program ma za zadanie wygrać z partnerem sparingowym,
- polegające na wiernym odtworzeniu funkcji binarnej, której pełna specyfikacja jest dana jako sformułowanie zadania.

Weryfikacja eksperymentalna w rozdziale 11 jest związana z minimalizacją błędu regresji funkcji ciągłej.



Wkład rozprawy w rozwój nauki stanowią rozdziały 8-12. Autor kompleksowo opracował istotny wycinek zagadnień związanych z wyborem funkcji celu na potrzeby realizacji uczenia za pomocą optymalizacji z użyciem programowania genetycznego. Wkład ten jest ważki i, co warto szczególnie podkreślić, został pozytywnie zweryfikowany przez Autora przed opracowaniem rozprawy. Wzmiankowane rozdziały stanowią opracowanie opublikowanych treści artykułów i referatów konferencyjnych, których współautorem jest Autor rozprawy.

Rozprawa ma charakter przeglądu ciekawych i oryginalnych pomysłów oraz ich weryfikacji eksperymentalnej z użyciem akademickich zadań testowych. Chwilami brakowało mi odniesienia do złożoności problemów świata rzeczywistego. Jednak rozległości tematyki rozprawy jest tak duża, że znacznie przekracza ona typowy dla rozpraw doktorskich zakres badanych pomysłów.

### **3. Ocena poprawności przedstawionych rozważań**

Treść rozprawy jest bardzo starannie przygotowana, z dużą dbałością zarówno o weryfikację eksperymentalną, jak i motywację teoretyczną. Autor opracowuje wyniki eksperymentów za pomocą testów statystycznych, co uwiarygadnia wnioskowanie o wzajemnej przewadze algorytmów. Zadania testowe użyte przez Autora stanowią „złoty standard” w przecięciu dziedzin programowania genetycznego i maszynowego uczenia w grach. Motywacja na gruncie obserwacji teoretycznych jest również dobrze udokumentowana. Rozważania są w mojej ocenie w pełni poprawne i dobrze udokumentowane.

### **4. Ocena jakości przeglądu stanu wiedzy zaprezentowanego w rozprawie**

Przegląd literatury w zakresie programowania genetycznego oraz wybranych zagadnień maszynowego uczenia, związanych z indukcją oprogramowania, świadczy o dużej erudycji Autora. Jedyny element, którego obecności nie byłem w stanie zrozumieć, stanowiły rozważania o koewolucji (rozdział 3), wprowadzający pojęcia, do których odwołanie nie wnosi wiele do treści rozdziałów 8-12.

Treść rozdziałów 1-7 wskazuje, że Autor ma zacięcie publicystyczno-dydaktyczne. Przykładowo, rozdział 6 jest bardzo zgrabnym esejem o wybranych aspektach ograniczeń wprowadzanych przez koncept funkcji celu. W wielu rozdziałach pojawiają się proste przykłady ilustrujące wprowadzane pojęcia. To sprawia, że rozprawę czyta się gładko.

Podsumowując, uważam że Autor wykazał się w rozprawie głębokim zanurzeniem w literaturze przedmiotu oraz dużym wyczuciem aktualnych kierunków badań.

### **5. Inne uwagi**

Rozprawa jest jedną z najlepszych, jakie miałem okazję recenzować. Autor zadbał o to, aby nie pozostawiać niedokończonych wątków, zadbał także o wzajemną spójność zakresu i sposobu badań eksperymentalnych, starannie umotywowował własne koncepcje i skomentował wyniki. Wybór języka angielskiego przez osobę nie władającą nim od urodzenia jest zawsze w mojej ocenie kontrowersyjny, ale szczęśliwie rozprawa daje się czytać bez zbytecznego dodatkowego wysiłku, chociaż nie udało się w niej uniknąć nielicznych błędów językowych.



## 6. Wniosek końcowy

Biorąc pod uwagę opinie zaprezentowane w poprzednich punktach i wymagania zdefiniowane przez artykuł 13 Ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym (z późniejszymi zmianami)<sup>1</sup>, a także wymagania stawiane rozprawom doktorskim przez obyczaj naukowy, moja ocena rozprawy pod względem trzech podstawowych kryteriów jest następująca:

A. Czy rozprawa zawiera oryginalne rozwiązanie problem naukowego?

<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zdecydowanie TAK	Raczej TAK	Trudno powiedzieć	Raczej NIE	Zdecydowanie NIE

B. Czy po przeczytaniu rozprawy zgadzasz się, że kandydat posiada ogólną wiedzę teoretyczną w dyscyplinie Informatyka?

<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zdecydowanie TAK	Raczej TAK	Trudno powiedzieć	Raczej NIE	Zdecydowanie NIE

C. Czy kandydat umiejętność samodzielnego prowadzenia pracy naukowej?

<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zdecydowanie TAK	Raczej TAK	Trudno powiedzieć	Raczej NIE	Zdecydowanie NIE

Wnoszę zatem o dopuszczenie Autora do dalszych etapów przewodu doktorskiego.

Jednocześnie wnoszę o uznanie rozprawy za wyróżniającą się, biorąc pod uwagę zarówno recenzowaną treść, jak i imponujący dorobek publikacyjny Doktoranta.



Jarosław Arabas

<sup>1</sup> [http://www.nauka.gov.pl/g2/oryginal/2013\\_05/b26ba540a5785d48bee41aec63403b2c.pdf](http://www.nauka.gov.pl/g2/oryginal/2013_05/b26ba540a5785d48bee41aec63403b2c.pdf)