



Kraków 2009-03-24

Recenzja rozprawy doktorskiej mgr inż. Stefana Kotowskiego pt. „Analiza algorytmów genetycznych jako układów dynamicznych”

1. Ogólna charakterystyka pracy

Przedłożona do recenzji rozprawa doktorska mgr inż. Stefana Kotowskiego pt. „*Analiza algorytmów genetycznych jako układów dynamicznych*” została zrealizowana pod kierunkiem Prof. dr hab. inż. Witolda Kosińskiego. Praca stanowi manuskrypt o łącznej objętości 111 stron zawierający 3 rysunki. Tekst pracy podzielono na 5 rozdziałów uzupełnionych nienumerowanym streszczeniem w języku angielskim. Praca zawiera bogaty, złożony ze 109 pozycji spis literatury.

Tezy rozprawy podane na stronie 8 brzmią następująco:

1. Istnieje rozkład graniczny i można go opisać jawną zależnością. Zależność ta wskazuje na sposób ulepszania algorytmu.
2. Dla każdego wyjściowego algorytmu genetycznego istnieje algorytm genetyczny, optymalny w sensie probabilistycznym.
3. Możliwa jest klasyfikacja algorytmów genetycznych na podstawie ich entropii i wymiaru fraktalnego trajektorii. Może być ona pożyteczna przy projektowaniu następnych algorytmów genetycznych.

Zadania, jakie zrealizował Autor w rozprawie można scharakteryzować następująco:

1. Przedstawił obszerne tło obejmujące opis wielu technik poszukiwania stochastycznego z grupy nazywanej algorytmami genetycznymi. W ich opisie ograniczył się do przypadków z dyskretną, skończoną przestrzenią poszukiwań. Opisał również dwie tradycyjne metody analizy algorytmów genetycznych takie jak tzw. twierdzenie o schematach, hipoteza bloków budujących oraz tzw. twierdzenie NFL (No Free Lunch). Całość tych rozważań zawarł on w

Rozdziale 1 rozprawy. W rozdziale tym zasygnalizowano również problematykę definiowania izomorfizmów służących do klasyfikacji algorytmów.

2. Formułuje podstawy modelowania rozważanej grupy algorytmów genetycznych przy pomocy łańcuchów Markowa bazujący na ideach i wynikach grupy Vose. W tej części zawarto ważne wyniki dotyczące klasyfikacji zachowania się algorytmów genetycznych wykorzystując jego zachowania asymptotyczne. W tej części znajduje się również najciekawszy bodajże wynik pracy dotyczący interpretacji hipotezy No Free Lunch dla rozważanej klasy algorytmów. Wyniki tej części zawarto w Rozdziale 3 rozprawy.
3. Kolejny Rozdział 4 zawiera drugą klasyfikację behawioralną algorytmów genetycznych opartą o wymiar fraktalny ich trajektorii. Zaproponowaną klasyfikację ilustruje kilka przykładów obliczeniowych.

2. Ocena pracy

Autor podjął w pracy trudne zagadnienie klasyfikacji behawioralnej algorytmów genetycznych pozwalającej na ich porównanie w zakresie skuteczności rozwiązania zadań optymalizacji globalnej. Problem ten, dotąd w ogólności nierozwiązany jest jak się wydaje najtrudniejszy w analizie asymptotycznej tej grupy algorytmów.

Istnieje wiele podejść pozwalających ocenić różne aspekty zachowania się algorytmów ze skończonym uniwersum genetycznym. Historycznie pierwsze pochodzące od Hollanda a później rozwinięte i częściowo doprecyzowane przez Whitleya znane pod nazwą twierdzenia o schematach pozwala na częściowe statystyczne oszacowanie (w sensie wartości oczekiwanej) wydajności eksploatacyjnej algorytmu (zdolności do uzyskiwania lepszego wyniku średniego w populacji) w pojedynczym kroku ewolucji. Wielu uznanych autorów zgodnie twierdzi, że wynik ten nie pozwala na asymptotyczne rozszerzenie, tj. uzyskanie pewnego oszacowania dla nieskończonej, a nawet dla skończonej liczby kroków poprzez jego iterowanie.

Najbardziej obiecującym jak się wydaje sposobem analizy asymptotycznej algorytmów genetycznych jest przedstawienie ich jako procesów stochastycznych, lub bardziej ogólnie jako układów dynamicznych w odpowiednio zdefiniowanej przestrzeni stanów. Z uwagi na niedeterministyczny charakter kolejnych iterat przestrzeń tę budują miary probabilistyczne na zbiorze wszystkich osiągalnych populacji. W rozważanym przypadku skończonego uniwersum kodów zbiór ten jest również skończony. Kluczowe wyniki tego podejścia uzyskane w grupie Vose pozwalają na sformułowanie warunku koniecznego dla istnienia miary niezmienniczej, będącej granicą ciągu miar próbkowania przy ilości epok zmierzającej do nieskończoności, oraz pokazują pewne regularne zachowanie tych miar w sytuacji, gdy rozważamy algorytmy operujące na coraz to większych populacjach. Wyniki te pozwalają sformułować jakościowe kryteria, przy których algorytm może rozwiązać stawiany przed nim problem, oraz gwarantują zbieżność trajektorii algorytmów o skończonych populacjach do trajektorii uzyskanej poprzez iteracje tzw. heurystyki, odwzorowania reprezentującego oczekiwane zachowanie się algorytmu w każdym kroku, czyli postępowania deterministycznego wykorzystującego maksimum informacji, jakie dać mogą populacje przetwarzane w ustalony sposób, tj. przy pomocy ustalonych, niezmiennych w czasie operatorów genetycznych.

Rozważania te, niestety nie podają sposobu klasyfikacji algorytmów w przypadku rozwiązywania różnych problemów związanych z różnymi funkcjami celu. Taka wyczerpująca (zupełna) klasyfikacja pozwoliłaby na odpowiedni dobór operatorów genetycznych i ich parametrów dla rozwiązania zadanego problemu przy najlepszej proporcji kosztów do ilości informacji uzyskanej o rozwiązaniu. Wiele faktów wskazuje na to, że problem takiej klasyfikacji jest dobrze postawiony i istotny. Jednym z takich wyników jest hipoteza No Free Lunch mówiąca, że w zadanej klasie algorytmów optymalizacji globalnej rozwiązującej zadany zbiór problemów dla każdego algorytmu istnieje problem, dla którego osiąga on najlepszą wydajność w całej klasie i jest również taki, dla którego jest on algorytmem najgorszym. Jakkolwiek bardzo ogólne uzasadnienia tej hipotezy wydaje się trudne, istnieje kilka przykładów jej precyzyjnej interpretacji dla konkretnych klas algorytmów i rozwiązywanych przez nie problemów (prace grupy Poli).

Najważniejszy moim zdaniem wynik Autora (Twierdzenie 3.4.8) pokazuje, że w klasie algorytmów genetycznych przeszukujących skończony zbiór, opisanych stacjonarnymi, ergodycznymi łańcuchami Markowa rozwiązujących dany problem optymalizacji globalnej istnieje algorytm „najlepszy”, który rozwiązuje ten problem w 1 kroku iteracji (jednej epoce). Przez rozwiązanie problemu Autor rozumie osiągnięcie przez algorytm odpowiedniej, niezmienniczej miary próbkowania. W tym samym rozdziale Autor podaje również inny ważny wniosek (Wniosek 3.4.1) mówiący o asymptotycznej równoważności prostego algorytmu genetycznego i przesunięcia Bernoulliego dających możliwość stosowania metod analizy układów dynamicznych do analizy algorytmów genetycznych.

Drugą ważną grupą wyników są efekty rozważań dotyczących asymptotycznych zachowań miar próbkowania dla populacji skończonych definiowanych jako asymptotyczna stabilność i punktowa asymptotyczna stabilność. Autor podaje precyzyjne definicje tych pojęć związanych z istnieniem odpowiednich miar niezmienniczych łańcucha Markowa modelującego algorytm genetyczny ze skończonym kodowaniem oraz ich charakteryzację w postaci twierdzeń 2.6.2 do 2.6.4 dla algorytmów z selekcją i mutacją. Wprowadzone pojęcia asymptotycznej stabilności oraz licznosc zbioru kodów (chromosomów) osiągalnych z dowolnego chromosomu w skończonej ilości kroków pozwala na zbudowanie innej klasyfikacji przedstawionej na końcu Rozdziału 2.

3. Uwagi krytyczne i dyskusyjne

Słabą stroną rozprawy jest redakcja manuskryptu. Zawiera on wiele fragmentów powtórzonych oraz wiele fragmentów sformułowanych niejasno. Nie jest również oczywista zależność pomiędzy podstawową częścią manuskryptu a Dodatkiem. Zawiera on brakujące w części podstawowej pracy dowody kluczowych twierdzeń. W innych rozdziałach dubluje lub pomija pewne fragmenty tekstu podstawowego (patrz sekcje 4.2 pracy i 3.2 Dodatku).

Pewnym mankamentem redakcyjnym jest również brak spisu oznaczeń i symboli, który w znacznym stopniu ułatwiłby czytanie rozprawy i uchronił Autora od licznych kolizji oznaczeń.

Jakkolwiek Autor definiuje precyzyjnie i analizuje dwa rodzaje zachowań asymptotycznych algorytmu genetycznego nazwane asymptotyczną stabilnością i

asymptotyczną punktową stabilnością nie wyjaśnia dokładnie, co rozumie przez rozwiązanie problemu optymalizacji globalnej.

Twierdzenie 2.6.2 podaje wkł. dla asymptotycznej, punktowej stabilności. Własność ta nie będzie zachodzić dla poprawnie zdefiniowanych prostych algorytmów genetycznych z selekcją i mutacją, tj. takich, dla których prawdopodobieństwo mutacji jest silnie większe od zera. W tym przypadku każdy chromosom może być osiągnięty w jednym kroku z dowolnego innego z dodatnim prawdopodobieństwem. Ta grupa algorytmów, dla których $Z^* = Z$ wpada w punkty 2 i 4 klasyfikacji. Istnienie i silna dodatniość miary niezmienniczej w tym przypadku wynika z twierdzenia ergodycznego (por. Vose, Nix 1992).

W Twierdzeniu 3.4.8 podano postać macierzy przejścia algorytmu najlepszego. Wielka szkoda, że Autorowi nie udało się zidentyfikować macierzy od parametrów operacji genetycznych. Jedyne wyniki idące w tym kierunku (sekcja 3.4.3) mówi o postaci macierzy granicznej dla algorytmu złożonego wyłącznie z selekcji elitarniej.

Wniosek 3.4.1 powinien zostać wyczerpująco uzasadniony w miejscu dokładnie określonym w pracy.

Sposób prowadzenia eksperymentów opisanych w rozdziale 4.3 jest niejasny i mało przekonujący. Nie wyjaśniono, co oznaczają konfiguracje od 1 do 12, oraz które punkty wykresu wiążą się z poszczególnymi funkcjami benchmarkowymi.

Dodatkowo spis usterek i pomyłek zauważonych w tekście rozprawy zamieszczono w postaci załącznika na końcu recenzji.

4. Wnioski końcowe

Podsumowując stwierdzam, że teza rozprawy została uzasadniona. Zawiera ona nowe spojrzenie na sposoby analizy wydajności algorytmów genetycznych z kodowaniem skończonym. Wysoko oceniam najważniejsze jej wyniki, jakimi są interpretacja hipotezy No Free Lunch dla algorytmów ze skończonym kodowaniem opisanych ergodycznym łańcuchem Markowa (Twierdzenie 3.4.8) wraz z propozycją klasyfikacji opartą o wymiar fraktalny trajektorii jak również analiza asymptotycznych zachowań miar próbkowania (asymptotycznej stabilności i punktowej asymptotycznej stabilności, Twierdzenia 2.6.1 – 2.6.4).

Praca spełnia wymagania stawiane rozprawom doktorskim w zakresie nauk technicznych w dyscyplinie informatyka i może być dopuszczona do publicznej obrony.



Załącznik – spis niektórych usterek zauważonych w pracy:

1. We wzorze zawartym w tekście na 7 stronie (4 strona dodatku) pojawia się kolizja związana z oznaczeniem $d_m^x(1)$.
2. W definicji na stronie 7 (str. 5 dodatku) pojawiają się nieznanne symbole d^x i X . Ponadto stosowana jest strzałka \mapsto w określeniu warunku dla funkcji a , używana zwykle dla oznaczenia funkcji częściowych, co jest w sprzeczności z poprzednią formułą.
3. Sekcja 1.4. Co Autor rozumie pod pojęciem zbieżności używanym w tym rozdziale?
4. Podany w pracy wzór (1.17) stanowiący wynik tzw. twierdzenia schematów jest nieprawdziwy. Jego krytykę znaleźć można w licznych pracach, np. w pracy Podsiadły opublikowanej na pierwszej konferencji KAEGiOG w 1996 roku.
5. Skąd Autor zaczerpnął definicję algebry abstrakcyjnej, którą używa w sekcji 1.6.2? Czy pomiędzy postulowanymi działaniami tej struktury zachodzą jakieś związki?
6. We wzorze (2.21) obie nierówności powinny być chyba ostre?

