

doc. dr hab. Janusz Szczepański
Instytut Podstawowych Problemów Techniki
Polska Akademia Nauk
ul. Świętokrzyska 21
00-049 Warszawa
e-mail: jszczepa@ippt.gov.pl

Warszawa, 18 marzec 2009 r.

Recenzja

Rozprawy doktorskiej mgr inż. Stefana Kotowskiego: „Analiza algorytmów genetycznych jako układów dynamicznych”.

Praca doktorska mgr. inż. Stefana Kotowskiego, o objętości 72 stron, składa się z pięciu Rozdziałów i spisu literatury obejmującej 101 pozycji. Za sugestią recenzentów autor uzupełnił rozprawę o Dodatek, w którym zawarł istotne szczegóły dowodów pewnych twierdzeń przedstawionych w pracy (związanych z warunkami punktowej asymptotycznej stabilności odpowiedniego modelu), naniósł kilka poprawek oraz podał precyzyjne definicje ważniejszych pojęć występujących w pracy. Dysertacja poświęcona jest, jak powiedziano w tytule, docelowo sformułowaniu i analizie algorytmów genetycznych w języku teorii układów dynamicznych. Algorytm genetyczny to rodzaj algorytmu przeszukującego przestrzeń potencjalnych rozwiązań problemu w celu wyszukania najlepszego (optymalnego) rozwiązania. Sposób działania algorytmów genetycznych przypomina zjawisko ewolucji biologicznej, gdyż właśnie z biologii czerpane były inspiracje przez twórców algorytmów. Wspólnymi cechami algorytmów ewolucyjnych, do których zaliczają się algorytmy genetyczne, odróżniającymi je od tradycyjnych metod optymalizacji, są: stosowanie operatorów genetycznych dostosowanych do postaci rozwiązań, równoległe przeszukiwanie przestrzeni rozwiązań z różnych punktów (obszarów), ukierunkowanie procesu przeszukiwania w oparciu o aktualne rozwiązania oraz wprowadzenie elementów losowych. Przebieg samego algorytmu jest następujący:

- losowana jest (według przyjętych reguł) pewna populacja początkowa, która poddawana jest ocenie w wyniku której najlepiej przystosowane osobniki biorą udział w procesie reprodukcji.
- chromosomy wybranych osobników są krzyżowane,
- zachodzi proces mutacji, czyli zmian losowych.

Rodzi się następne pokolenie i algorytm powraca, o ile nie znaleziono dostatecznie dobrego rozwiązania do wyjściowego/początkowego kroku.

Z kolei przez układ dynamiczny (z dyskretnym „czasem”) rozumie się deterministyczne odwzorowanie (operator ewolucji) działający na pewnej przestrzeni z miarą. Odwzorowanie to musi zachowywać tę miarę (innymi słowy miara ta jest niezmiennicza). Układy takie modelują wiele zjawisk fizycznych, systemów biologicznych oraz ekonomicznych. Na przestrzeni kilkudziesięciu lat teoria układów dynamicznych była intensywnie rozwijana, uzyskano szereg silnych twierdzeń, zwłaszcza związanych z tzw. ergodycznymi własnościami takich układów.

W recenzowanej rozprawie autor formułuje, ewolucję zadaną przez algorytmy genetyczne, w języku procesów stochastycznych (jednorodne łańcuchy Markowa), a następnie wykorzystuje do analizy algorytmów genetycznych, związki pomiędzy (stacjonarnymi) procesami stochastycznymi, a układami dynamicznymi. Głównym celem pracy jest zatem wykorzystanie silnego aparatu układów dynamicznych do analizy (w szczególności klasyfikacji) algorytmów genetycznych.

Jeśli chodzi o strukturę pracy to autor kolejno przedstawia:

W Rozdziale I przypomniana została historia algorytmów genetycznych, opisane zostały główne typy tych algorytmów. Ponadto, w sposób przejrzysty i dobrze sformułowany od strony matematycznej określono podstawowe obiekty występujące w teorii algorytmów genetycznych. Na zakończenie, autor szczególną uwagę poświęca tzw. problemowi *No Free Lunch* tzn. nieistnieniu uniwersalnego algorytmu optymalizacyjnego. Rozdział ten kończy się przypomnieniem pojęć izomorfizmu, które to pojęcia odgrywają później istotną rolę w klasyfikacji algorytmów genetycznych. Autor najpierw przypomina definicję izomorfizmu „algebraicznego” (zachowana jest struktura algebraiczna), a następnie

nawiązuje do grup przekształceń zachowujących miarę oraz izomorfizmów pomiędzy takimi obiektami. Istotną własnością dla izomorfizmów są ich niezmienniki, w szczególności dla izomorfizmów metrycznych takim niezmiennikiem jest entropia.

W Rozdziale II autor formułuje matematyczne podstawy algorytmów genetycznych w kontekście probabilistycznym (koncentrując się na SGA - Simple Genetic Algorithm); definiuje najpierw zbiór możliwych populacji (kodowanie - binarna reprezentacja osobników), a następnie w języku probabilistycznym operator selekcji (zależny od funkcji przystosowania), operator mutacji (zadany przez macierz prawdopodobieństw mutacji) oraz operator krzyżowania (zdefiniowany za pomocą macierzy określających prawdopodobieństwa krzyżowania). Pozwala to na zdefiniowanie tzw. macierzy przejścia T , a tym samym na probabilistyczny opis ewolucji algorytmu genetycznego jako procesu Markowa (zatem, w tym momencie jest to jeszcze proces stochastyczny), w którym stanami są składowe korespondujące do rozkładu prawdopodobieństwa zdefiniowanego przez „wektor populacji” (podrozdział 2.2, str. 32). Istotną rolę w zrozumieniu tego podejścia odgrywa tutaj formuła 1.5 określająca prawdopodobieństwa warunkowe. Zakładając jednorodność tak określonego procesu Markowa zasadniczym pytaniem dla tego typu macierzy jest istnienie rozkładów stacjonarnych. Szczególnie istotny przy założeniu jednorodności wydaje się być odpowiedni dobór jeśli operatora selekcji.

W paragrafie 2.6 sformułowane zostały także ważne twierdzenia (2.6.2, str. 39) dotyczące punktowej asymptotycznej stabilności dla modeli z selekcją i mutacją dla populacji skończonych. Dowody tych twierdzeń zostały podane właśnie w załączonym Dodatku.

W Rozdziale III autor przedstawia w jaki sposób otrzymany łańcuch Markowa może być traktowany jako układ dynamiczny. W istocie, chodzi tutaj o fakt, że w przypadku granicznym łańcuch taki jest izomorficzny z konkretnym procesem stochastycznym, tzw. schematem Bernoulliego. Jeśli spojrzeć na schemat Bernoulliego (opisany z kolei w podrozdziale 4.1.1) w ten sposób, że przestrzenią jest zbiór ciągów, na którym to zbiorze określona jest przy pomocy zbiorów cylindrycznych miara (wykorzystuje się tutaj twierdzenie Kołmogorowa o rozkładach zgodnych) i na tej przestrzeni działa operator przesunięcia zachowujący tę miarę, to mamy dobrze określony układ dynamiczny w

sensie, o którym było wspomniane na początku recenzji (tzw. przesunięcie Bernoulliego). Z kolei dla przesunięcia Bernoulliego automatycznie mamy analitycznie wyliczoną wartość entropii (precyzyjniej tempa entropii, ang. entropy rate) oraz później wykorzystywane kluczowe Twierdzenie Ornsteina.

W Rozdziale III przedstawione są też główne twierdzenia dotyczące rozkładu granicznego dla algorytmu genetycznego z selekcją elitarną (Paragraf 3.4.3). Ponadto opisana została także postać operatora granicznego. Z własności tej autor wywnioskował postać algorytmu (proponowana graniczna macierz ma wszystkie wiersze identyczne równe rozkładowi stacjonarnemu), który w pierwszym kroku daje największe prawdopodobieństwo wyznaczenia rozwiązania optymalnego. Co więcej, autor zasugerował (Wniosek 3.4.3 str. 50) sposób wyznaczania współczynników macierzy granicznej/optymalnej wykorzystując proporcje czasów spędzanych przez pojedynczą, typową trajektorię procesu w poszczególnych stanach (uśrednianie po długich czasach). Milcząco korzysta się tutaj z twierdzeń ergodycznych.

W Rozdziale IV autor przypomina podstawowe wielkości charakteryzujące układy dynamiczne i procesy stochastyczne, mianowicie pojęcie entropii, a także wymiaru fraktalnego. Ponadto w Dodatku autor załączył precyzyjne matematycznie definicje tych koncepcji. W kontekście entropii kluczowe jest wspomniane twierdzenie Ornsteina, które mówi, że dwa dowolne przesunięcia Bernoulliego posiadające tą samą entropię są izomorficzne. Umożliwia to klasyfikację takich układów. Następnie autor przypomina twierdzenie Choe (4.2.2, str. 58), które dla przesunięcia Bernoulliego wiąże wymiar Hausdorffa z entropią. Zatem chcąc ocenić entropię można to uczynić poprzez oszacowanie wymiaru Hausdorffa.

Nawiązując do tej własności w podrozdziale 4.3 przedstawione zostały ciekawe wyniki badań eksperymentalnych (symulacji numerycznych). Napisano oprogramowanie (w języku Java) do realizacji algorytmów wyznaczania właśnie wymiarów fraktalnych trajektorii. Oprogramowanie ma bardzo elastyczną strukturę, składa się z kilku modułów/interfejsów do tworzenia nowych klas: Chromosom, Estymator, Selector, Mutator, Reproduktor, Improver.

Podstawowy moduł systemu wylicza wymiar pudełkowy dowolnego zbioru punktów w n -wymiarowej przestrzeni. Zaobserwowano ciekawy fakt potwierdzający przyjęte w pracy wnioski, mianowicie dla tej samej funkcji testowej oraz konfiguracji wymiar pudełkowy trajektorii wyraźnie skupia się wokół pewnych wartości.

Problemy, które wymagają dalszej analizy i/lub wyjaśnienia

Głębszej analizy wymaga problem, w jaki sposób entropia otrzymanego układu granicznego zależy od „parametrów” algorytmu genetycznego, w szczególności od prawdopodobieństw mutacji i od prawdopodobieństw selekcji (formuła 2.5 str. 32) oraz przyjętych wartości prawdopodobieństw w macierzach dla operatora krzyżowania. Należy wziąć pod uwagę, że jeśli prawdopodobieństwa mutacji są bardzo duże (bliskie jeden) to entropia jest wtedy jednak mała. Wiadomo, że przysunięcia Bernoulliego ma największą entropię gdy bity/symbole są generowane z równym prawdopodobieństwem. Z kolei łańcuch Markowa, w którym z dużym prawdopodobieństwem „przeskakuje” się do innego stanu też mają małą entropię.

W Dodatku na stronie 22 nie jest jasne stwierdzenie:

„Kwestia optimum może być niejednoznaczna, gdyż dwa algorytmy dające rozkład graniczny, w jednym kroku mogą mieć różny rozkład graniczny. Wówczas lepszy jest ten, który daje większe prawdopodobieństwo najlepszemu rozwiązaniu. Należy więc uwzględnić entropię rozkładu granicznego.” Jak się ma to twierdzenie, do faktu że dla stacjonarnych procesów Markowa (o skończonej liczbie stanów) jeśli ograniczymy się do składowych ergodycznych, to na poszczególnych składowych mamy jednoznaczny rozkład graniczny.

Wydaje mi się, że bardziej adekwatnym tytułem rozprawy byłby tytuł nawiązujący do modelowania algorytmów genetycznych łańcuchami Markowa.

Proponowany sposób modelowania zakłada pewne ograniczenia, mianowicie zakłada, że w obrębie składowych „wektorów populacji” ewolucja dla każdego takiego chromosomu

(takich samych chromosomów w obrębie populacji) odbywa się w identyczny sposób. Jest to pewne ograniczenie rozważanych algorytmów.

Centralnym wynikiem eksponowanym w dysertacji jest określenie w oparciu o istnienie rozkładu granicznego dla zadanej macierzy Markowa, [która zależy od własności superpozycji macierzy selekcji, macierzy mutacji i macierzy krzyżowania] postaci optymalnego operatora (w sensie probabilistycznym), który w jednym kroku z dowolnego rozkładu początkowego generuje rozkład graniczny/optimalny. Przytoczone w rozprawie twierdzenia (na przykład Tw. 3.3.3) dotyczą sytuacji gdy proces Markowa jest jednorodny. Nasuwa się zatem pytanie, jakie precyzyjnie warunki muszą być nałożone na operatory ewolucji (selekcji, krzyżowania, mutacji) żeby można było założyć jednorodność w czasie indukowanego procesu Markowa? Częściowa dyskusja tego problemu jest w podrozdziale 3.3.4.

Warto zauważyć ponadto, że ważne Twierdzenie 3.4.8 (Dodatek, strona 21) w rozprawie nie rozstrzyga, czy graniczny operator (otrzymany z rozkładu stacjonarnego) dla algorytmu genetycznego jest również algorytmem genetycznym.

Dobrze byłoby sformułować precyzyjnie, kiedy w przypadku macierzy stochastycznej wartości własnej równej 1 odpowiada przestrzeń niezmiennicza jedno-wymiarowa. Oznacza to, że dla takiego procesu Markowa istnieje dokładnie jeden rozkład graniczny. Wiąże się to z ergodycznością procesu.

Ciekawa byłaby głębsza biologiczna interpretacja otrzymanych wyników.

Główne wyniki rozprawy

Za główne wyniki rozprawy można uznać:

Przejrzyste i precyzyjne sformułowanie związku pomiędzy ewolucyjnymi algorytmami genetycznymi, a łańcuchami Markowa, a następnie układami dynamicznymi.

Zaproponowanie klasyfikacji algorytmów genetycznych za pomocą wielkości charakteryzujących graniczne dla nich (via łańcuchy Markowa) układy dynamiczne (przesunięcie Bernoulliego). Wielkościami takimi są (str. 63): entropia, wymiar fraktalny (Hausdorffa, pudełkowy, informacyjny). Zwraca uwagę przygotowanie wspomnianej implementacji do estymowania wymiaru Hausdorffa.

Jeśli chodzi o twierdzenie *NFL* to ważnym wynikiem rozprawy jest zaobserwowanie, a właściwie potwierdzenie, że optymalny algorytm istotnie zależy od własności (parametrów) problemu, który mamy rozwiązać.

Na stronie 58 autor konkluduje, że analizowana entropia otrzymanego procesu Markowa jest funkcją prawdopodobieństw mutacji i selekcji. Ciekawa wydaje się obserwacja/wynik, że entropia wzrasta ze wzrostem prawdopodobieństwa mutacji i maleje gdy rośnie „nacisk selekcyjny”.

Za główne Twierdzenia rozprawy można uznać Twierdzenia 3.3.2, 3.3.3 oraz 3.3.4, które mówią, że algorytmy genetyczne, w których parametry są dostrajane (poprzez dołączenie parametrów do chromosomu i poddaniu ich działaniu operatorów genetycznych razem z całym chromosomem lub poprzez dostrajane tych parametrów w zależności od wartości funkcji przystosowania lub jej rozkładu) są asymptotycznie stabilne i ich rozkład graniczny nie zależy od populacji początkowej.

Ponadto, ważne i ciekawe są sformułowania (i dowody w Dodatku) Twierdzeń dla modeli z selekcją i mutacją dla **populacji skończonych** (Tw. 2.6.2, 2.6.3, 2.6.4).

Rozprawa jest inspirująca i stanowi bardzo dobry materiał do dalszych analiz tych ważnych i złożonych problemów (optymalizacja).

Dorobek naukowy

Jeśli chodzi o dorobek naukowy to mgr inż. Stefan Kotowski, zgodnie z załączonym wykazem, jest autorem 11 prac naukowych w tym:

2 prace ukazały się w Archiwum Mechaniki,
3 artykuły w pismach o charakterze lokalnym i
6 prac w materiałach konferencyjnych.

Jest to dorobek, który jest w sposób istotny wystarczający dla osób ubiegających się o tytuł doktora.

Na szczególną uwagę zasługuje także fakt, że w roku 1959 mgr inż. Stefan Kotowski zdobył wyróżnienie na IX Olimpiadzie Matematycznej.

Zestawienie zauważonych usterek redakcyjnych

Praca zawiera szereg usterek redakcyjnych. Kilka z nich:

Str. 7

Wstęp. Definicja. Algorytm optymalizacyjny α jest przekształceniem określonym na D o wartościach w X

$$\alpha : D \rightarrow X$$

spełniającym dodatkowy warunek

$$\alpha : D \ni d \mapsto \{x \mid x \notin d^x\}$$

jest niedokładny. Prawa strona nie jest elementem X .

Str. 46

W zdaniu „W przypadku, gdy parametry mutacji, krzyżowania lub selekcji zależą od danej populacji i są dodatnie, rozkład prawdopodobieństw przejścia jest stały i niezależny od kroku macierzy i tworzy macierz Markowa” powinno być „nie zależą”.

Str. 58

W formule (4.2) jest błąd drukarski. Zamiast ε powinna być δ . Brakuje także nawiasu.

Brak jest czasami referencji do twierdzeń, z których autor korzysta (uzupełnione w Dodatku). Na przykład:

Str. 47 Twierdzenie 3.4.3

Str. 48 Twierdzenie 3.4.4

Autor nie ujednolicił przyjętego działania macierzy stochastycznej (macierzy przejścia) na wektor prawdopodobieństwa. Na ogół jest notacja prawostronna, ale czasami jest notacja lewostronna.

W paragrafie 4.1.1 są pewne nieścisłości koncepcyjne. Bardzo ważną rolę w rozważaniach w dysertacji odgrywa przesunięcie Bernoulliego. Jeśli chodzi o entropię (służącą autorowi do klasyfikacji) to jest ona liczona dla miary niezmienniczej, zatem powinno się rozpatrywać raczej ciągi dwustronne. W pracy nie jest jasno napisane, że chodzi o takie układy.

Podsumowanie

Oceniając oryginalne wyniki uzyskane (i wymienione powyżej) w rozprawie (oraz dorobek naukowy mgr inż. Stefana Kotowskiego) uważam, że recenzowana praca odpowiada warunkom określonym w Art. 13 Ustawy o stopniach naukowych i tytułach naukowych i stawiam wniosek o dopuszczenie jej do publicznej obrony.

Janina Kucperawka