

Piotr KOŁODZIEJCZYK, Kazimierz WANAT,
Andrzej JAWORSKI

ALGORYTMY GENETYCZNE W ZAGADNIENIACH GÓRNICZYCH

Część I

Streszczenie. Algorytmy genetyczne są to algorytmy wykorzystujące mechanizmy doboru naturalnego oraz zasady dziedziczności. Oparte na bazie ewolucyjnej formuły dziedziczenia najlepszych cech osobników najlepiej przystosowanych oraz losową, choć ukierunkowaną wymianą informacji, stanowią stosunkowo nową metodę poszukiwania właściwego rozwiązania. Algorytmy te efektywnie wykorzystują wcześniej uzyskane informacje do określenia nowego obszaru poszukiwań o spodziewanym wyższym stopniu adaptacji.

GENETIC ALGORITHMS IN THE MINING PROBLEMS. Part one

Summary. Genetic algorithms are those, which make use of the rules of natural choice and principles of inheritance. Based on the evolutionary principle of inheriting the best features of the best adapted individuals and random, but directed exchange of information constitute a comparatively new solution method. These algorithms effectively use previous information to define a new area of research of an expected higher level of adaptation.

1. Wprowadzenie

Algorytmy genetyczne są to algorytmy wykorzystujące mechanizmy doboru naturalnego oraz zasady dziedziczności. Oparte na bazie ewolucyjnej formuły dziedziczenia najlepszych cech osobników najlepiej przystosowanych oraz losową, choć ukierunkowaną wymianą informacji,

stanowią stosunkowo nową metodę poszukiwania właściwego rozwiązania. Algorytmy te efektywnie wykorzystują wcześniej uzyskane informacje do określenia nowego obszaru poszukiwań o spodziewanym wyższym stopniu adaptacji. Oznacza to, że w każdym nowym pokoleniu zostaje utworzona grupa osobników (ciągów bitowych) wynikająca z połączenia fragmentów najlepiej przystosowanych przedstawicieli poprzedniego pokolenia i określonej probabilistycznie nowej części składowej. Algorytmy genetyczne jako nowe narzędzie w zagadnieniach poszukiwania, optymalizacji i sterowania zrodziły się na bazie teorii systemów adaptacyjnych. Zostały rozwinięte przez Johna H. Hollanda oraz jego współpracowników z Uniwersytetu Michigan. Natomiast inspiracją do napisania tego artykułu była książka Davida Goldberga pt. "Gentic Algorithms in Search, Optimization, and Machine Learning", w tłumaczeniu na język polski "Algorytmy genetyczne i ich zastosowania".

2. Metody poszukiwań i optymalizacji

W ciągu kilkunastu ostatnich lat algorytmy genetyczne znalazły się w szerokim kręgu zainteresowań różnych dyscyplin naukowych, w tym biologii, biochemii, immunologii, informatyki, ekonomii i zagadnień inżynierskich. Przyczyną popularności i rosnącej liczby zastosowań algorytmów genetycznych jest ich podstawowa cecha - *odporność*, związek pomiędzy efektywnością i skutecznością w szerokim zakresie rozwiązywanych problemów oraz w różnych dziedzinach zastosowań. Są one również pozbawione wielu ograniczeń nakładanych na funkcję w przestrzeni poszukiwań, takich jak: istnienie pochodnych, ciągłość, jednomodalność itp. Dotychczas stosowane metody wyznaczania ekstremum, opisywane w literaturze przedmiotu, wyróżniały trzy główne trendy: poszukiwanie analityczne, przeglądowe (enumeratywne) i losowe.

Metody analityczne zasadniczo można podzielić na dwa rodzaje: metody bezpośrednie i metody pośrednie. Pierwsze z nich polegają na poszukiwaniu lokalnego maksimum (minimum) poprzez badanie funkcji w kierunku wyznaczonym przez lokalny gradient. Drugie dotyczą poszukiwania lokalnych ekstremów poprzez rozwiązanie układu równań (najczęściej nieliniowego) wynikającego z przyrównania gradientu funkcji celu do zera.

Obie metody te pomimo wielu różnych modyfikacji i ulepszeń wykazują brak odporności, ponieważ po pierwsze szukają optymalnego rozwiązania w najbliższym sąsiedztwie punktu startowego, czyli charakteryzują się lokalnym zakresem zbieżności, a po drugie ich zastosowanie wymaga istnienia pochodnych w całej przestrzeni rozwiązań dopuszczalnych.

Metody przeglądowe, jak sama nazwa wskazuje, przeglądają po kolei wszystkie punkty w skończonej przestrzeni poszukiwań, a algorytm każdorazowo oblicza wartości funkcji celu. Sposób ten choć atrakcyjny z uwagi na prostotę i intuicyjny charakter rozwiązywania problemu również wykazuje brak odporności z jednego tylko powodu - nieefektywności. Nawet w praktycznych zadaniach o umiarkowanym rozmiarze i złożoności sprawdzenie wszystkich możliwych punktów rozwiązań w użytecznym czasie jest nieosiągalne (pomimo rosnącej mocy obliczeniowej i szybkości przetwarzania współczesnych komputerów).

Metody poszukiwania losowego wykorzystujące błądzenie przypadkowe i inne schematy losowe oparte na szukaniu i zapamiętywaniu najlepszego rozwiązania pozbawione są wprawdzie ograniczeń metod analitycznych, lecz pod względem efektywności niewiele odbiegają od metod enumeratywnych. Niemniej idea wyboru losowego, zastosowana nie jako metoda rozwiązania, ale jako drogowskaz w ukierunkowanym procesie poszukiwań w zakodowanej przestrzeni rozwiązań, stanowi podstawową przesłankę algorytmów genetycznych.

3. Algorytmy genetyczne

Algorytmy genetyczne (AG) charakteryzują się czterema podstawowymi cechami:

- nie przetwarzają bezpośrednio parametrów zadania, lecz ich zakodowaną postać,
- prowadzą poszukiwania, wychodząc nie z pojedynczego punktu, lecz z pewnej ich populacji,
- korzystają tylko z funkcji celu, nie zaś z jej pochodnych lub innych pomocniczych informacji,
- stosują probabilistyczne, a nie deterministyczne reguły wyboru (D.E. Goldberg - 1995).

Mechanizm działania elementarnego algorytmu genetycznego jest stosunkowo prosty, gdyż wykorzystuje on zasadniczo trzy podstawowe operacje (procedury): reprodukcji, krzyżowania i mutacji.

Reprodukcja jest procesem, w którym indywidualne ciągi kodowe (osobniki) zostają powielone w stosunku zależnym od wartości, jakie przybiera dla nich funkcja celu (w biologii nazywana jest funkcją przystosowania). Reprodukcja zależna od przystosowania polega na tym, że osobniki o wyższym wskaźniku przystosowania mają większe prawdopodobieństwo wprowadzenia jednego lub kilku potomków do następnego pokolenia (generacji). W naturalnych populacjach operacja ta stanowi odzwierciedlenie przyrodniczej zasady przeżycia osobników najlepiej przystosowanych, gdzie wskaźnik przystosowania jest miarą zdolności osobnika do przetrwania i wydania potomstwa.

Algorytmicznie procedurę reprodukcji realizuje się poprzez symulację obrotowej tarczy, gdzie każdemu osobnikowi z populacji przyporządkowany jest jej wycinek proporcjonalny do wskaźnika przystosowania danego osobnika. Tworząc pokolenie potomne, proces uruchomienia tarczy wybiera jednego z kandydatów do reprodukcji. Dzięki temu osobniki lepiej przystosowane wprowadzają większą liczbę potomków do puli reproduktorów.

Po zakończeniu reprodukcji dokonuje się operacji *krzyżowania*, która przebiega w dwóch fazach. W pierwszej fazie następuje losowe kojarzenie osobników z puli reproduktorów w pary. Natomiast w fazie drugiej każda para poddawana jest procesowi krzyżowania właściwego, co odpowiada wzajemnej wymianie bitów w obu ciągach kodowych od losowo wybranego punktu krzyżowania i wytworzeniu w ten sposób dwóch nowych osobników do następnej generacji.

Ostatnim ogniwem podstawowego algorytmu genetycznego jest proces *mutacji*, który polega na sporadycznej (z niewielkim prawdopodobieństwem), przypadkowej zamianie elementu ciągu kodowego. W przypadku osobnika o kodzie dwójkowym oznacza to zamianę jedynki na zero i odwrotnie. Mutacja tak rozumiana jest więc błędzeniem przypadkowym w przestrzeni ciągów kodowych (osobników). Stosowana oszczędnie i traktowana jako mechanizm drugorzędny w stosunku do pierwotnego procesu reprodukcji i krzyżowania zabezpiecza niejako następną generację przed dominacją superosobników.

Wymienione powyżej cztery cechy - kodowanie parametrów, działanie na populacji, korzystanie z minimum informacji o zadaniu oraz zrandomizowane operacje reprodukcji, krzyżowania i mutacji składają się w efekcie końcowym na odporność algorytmu genetycznego oraz wynikającą stąd jego wyższość nad innymi, powszechnie stosowanymi technikami.

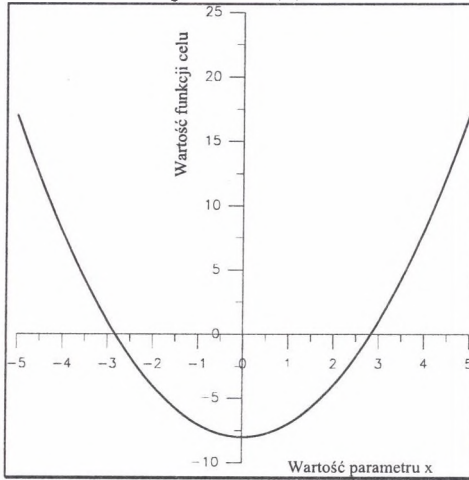
4. Wyniki symulacji komputerowych

Zbadanie możliwości zastosowania algorytmu genetycznego na potrzeby rozwiązywania zagadnień geofizyki górniczej, w tym lokalizacji ognisk wstrząsów sejsmicznych występujących w polskich kopalniach węgla kamiennego, rud miedzi, jak również węgla brunatnego stanowiło podstawową przesłankę zainteresowania się tą techniką rozwiązywania złożonych problemów numerycznych.. Ponieważ problem lokalizacji hipocentrum zjawiska sejsmicznego jest klasycznym zagadnieniem odwrotnym rozwiązywanym w skomplikowanej przestrzeni cztero- bądź trójwymiarowej, gdzie funkcja celu posiada minima lokalne, efektywne znalezienie minimum globalnego za pomocą algorytmu genetycznego może być przyczynkiem do zastosowań (AG) w innych obszarach działalności inżynierskiej.

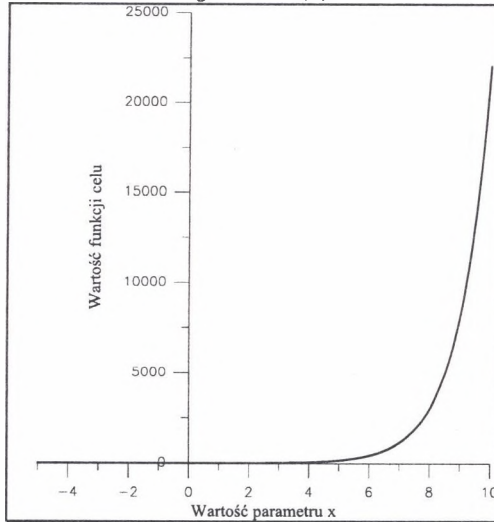
W dalszej części niniejszego artykułu przedstawiono wyniki działania algorytmu genetycznego w zadaniu poszukiwania minimum globalnego czterech wybranych typów funkcji jednej zmiennej :

- a. funkcja $f(x) = x^2 - 8$ w przedziale $-5 \leq x \leq 5$.
- b. funkcja $f(x) = e^x$ w przedziale $-5 \leq x \leq 10$.
- c. funkcja $f(x) = x^4 - x^3 - 24x^2 + 4x + 80$ w przedziale $-5 \leq x \leq 6$.
- d. funkcja $f(x) = \sin(x)\cos(15x)$ w przedziale $0 \leq x \leq \pi$.

W każdym z wyżej wymienionym zadaniu pokolenie (generację) reprezentowało 30 indywidualnych ciągów kodowych (osobników). Natomiast każdy ciąg kodowy (osobnik) charakteryzował zakodowaną wartość poszukiwanego parametru x w postaci 10-bitowego kodu binarnego.

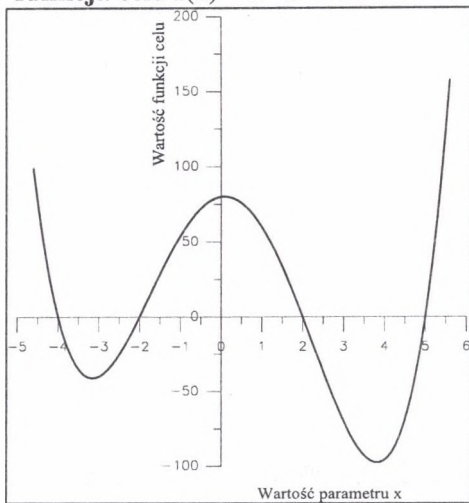
funkcja celu $f(x) = x^2 - 8$ 

	kod binarny	wartość parametr x	minimalna wartość funkcji celu
Pokolenie 0	0111110110	-0.09286	-7.99138
Pokolenie 16	0111111111	-0.00489	-7.99998

funkcja celu $f(x) = e^x$ 

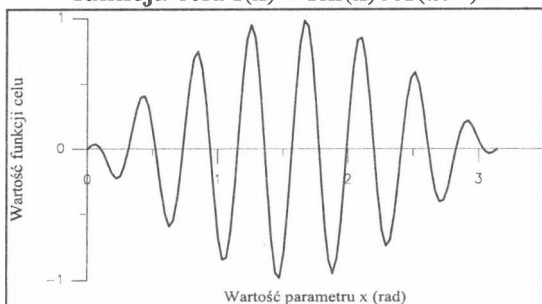
	kod binarny	wartość parametr x	minimalna wartość funkcji celu
Pokolenie 0	0000010111	-4.66276	0.00944
Pokolenie 6	0000000010	-4.97067	0.00694

funkcja celu $f(x) = x^4 - x^3 - 24x^2 + 4x + 80$



	kod binarny	wartość parametru x	minimalna wartość funkcji celu
Pokolenie 0	1101100000	3.84575	-97.71191
Pokolenie 6	1101011110	3.82620	-97.74145

funkcja celu $f(x) = \sin(x)\cos(15x)$



	kod binarny	wartość parametru x	minimalna wartość funkcji celu
Pokolenie 0	0111100110	1.49249	-0.91972
Pokolenie 9	0111011110	1.46792	-0.99433

Literatura

1. Dawid E. Goldberg: Algorytmy genetyczne i ich zastosowanie. WNT, Warszawa 1995.

Recenzent: Prof. dr hab. inż. Józef Dubiński

Wpłynęło do Redakcji 12.10.1997 r.

Abstract

Genetic algorithms are those, which make use of the rules of natural choice and principles of inheritance. Based on the evolutionary principle of inheriting the best features of the best adapted individuals and random, but directed exchange of information constitute a comparatively new solution method. These algorithms effectively use previous information to define a new area of research of an expected higher level of adaptation. It means that each new generation is made up by a group of individuals (binary sequences) effected by the elements of the best adapted representatives of the previous generation and the new constituent part defined by way of probability. Recently, genetic algorithms have been widely taken interest in by various sciences, including such as biochemistry, immunology, informatics, economy and many engineering branches.