

ALGORYTMY GENETYCZNE

opracowała
Katarzyna Kolonko

Matematyczne Koło Naukowe
Wydział Matematyczno-Fizyczny
Politechnika Śląska

I Wprowadzenie

Rozwój algorytmów genetycznych (Genetic Algorithms) został zapoczątkowany w roku 1975 pracą *Adaptation in Natural* Johna Hollanda z Uniwersytetu Michigan. Obecnie zastosowanie algorytmów genetycznych jest imponujące, stosowane są, bowiem one w szeregowaniu zadań, modelowaniu finansowym, optymalizacji funkcji oraz harmonogramowaniu. Gwałtownie przeżywany rozwój algorytmów genetycznych pobudził w ostatnim czasie różne obszary nauki i techniki min. fizykę, biologię, programowanie układów VLSI, a także zarządzanie. Rozróżnia się trzy zasadnicze grupy zastosowań AG: algorytmy przeszukujące (Search), optymalizujące (Optimization) i uczące (Learning). Algorytmy genetyczne wykorzystują poszukiwania oparte na mechanizmach doboru naturalnego oraz dziedziczności.

II Pojęcia

Algorytmy genetyczne stosują określenia wykorzystywane w genetyce.

Allel jest wartością danego genu, określaną także jako wartość cechy

Chromosom (ciąg kodowy) jest zbiorem osobników o określonej liczebności

Fenotyp jest zestawem wartości odpowiadających danemu genotypowi;

Gen (znak, detektor) jest zespołem chromosomów danego osobnika

Genotyp (struktura) jest to zespół chromosomów danego osobnika

Locus (pozycja) wskazuje miejsce położenia danego genu w chromosomie

Osobnikami populacji w algorytmach genetycznych są zakodowane w postaci chromosomów zbiory parametrów zadania (rozwiązania, punkty przestrzeni poszukiwań);

Populacja jest zbiorem osobników o określonej liczebności.

III Klasyczny Algorytm Genetyczny

Mechanizm działania klasycznego (elementarnego, prostego) algorytmu genetycznego opiera się na kopiowaniu ciągów i wymianie podciągów.

Na działanie Algorytmu genetycznego składają się cztery kroki:

(1) *Inicjacja* polega na losowym wyborze żądanej liczby chromosomów reprezentowanych przez ciągi binarne o określonej długości.

(2) *Ocena przystosowania* polega na obliczeniu funkcji przystosowania dla każdego chromosomu.

Im większa jest wartość tej funkcji tym dany chromosom jest lepiej przystosowany.

(3) *Selekcja* polega na wyborze chromosomów, które będą brały udział w tworzeniu nowej populacji

Selekcja to proces, w którym indywidualne ciągi kodowe zostają powielone w stosunku zależnym od wartości, jakie przybiera funkcja celu (przystosowania). Intuicyjnie funkcja f stanowi zysk, który chcemy zmaksymalizować. Reprodukacja polega na tym, że ciągi o wyższym przystosowaniu mają większe prawdopodobieństwo wprowadzenia jednego lub więcej potomków do następnego pokolenia. W naturalnych populacjach przystosowanie jest

miarą zdolności osobnika do uniknięcia drapieżnika, chorób i innych przeszkód na drodze do osiągnięcia dojrzałości i wydaniu potomstwa.

Reprodukcja następuje zgodnie z zasadą ruletki. Każdemu osobnikowi z populacji odpowiada sektor koła o rozmiarze proporcjonalnym do wartości funkcji przystosowania.

Następnie losujemy fragment koła (liczbę na ruletce), tyle razy, ile jest osobników w populacji. Osobniki, którym przyporządkowany jest większy wycinek koła, mają podwyższone szanse na przejście do następnego pokolenia.

W selekcji turniejowej dzieli się osobniki na podgrupy, i z nich wybiera się osobnika o najlepszym przystosowaniu. Rozróżnia się dwa wybory: deterministyczny i losowy.

W przypadku deterministycznym wyboru dokonuje się z prawdopodobieństwem równym 1, a w przypadku wyboru losowego z prawdopodobieństwem mniejszym od 1. Podgrupy mogą być dowolnego rozmiaru. Najczęściej dzieli się populację na podgrupy składające się z 2 lub 3 osobników.

W selekcji rankingowej osobniki populacji są ustawiane kolejno w zależności od wartości ich funkcji przystosowania. Każdemu osobnikowi przypisana jest liczba określająca jego kolejność na liście- randze. Liczba kopii każdego osobnika, wprowadzonych do populacji rodzicielskiej jest ustalana zgodnie z wcześniej zdefiniowaną funkcją, zależną od rangi osobnika.

(4) *Operatory genetyczne*. W klasycznym algorytmie genetycznym wyróżnia się operator krzyżowania oraz mutacji.

Krzyżowanie (crossover) jest wymiana fragmentu genotypu. Krzyżowanie proste w pierwszej fazie polega na kojarzeniu w sposób losowy ciągów z puli rodzicielskiej w pary, a w drugiej następuje proces krzyżowania. W sposób losowy z jednakowym prawdopodobieństwem (przyjmuje się $0.5 \leq p_c \leq 1$) wybierany jest punkt krzyżowania k spośród $w-1$ pozycji. A następnie zamieniane są wszystkie znaki od pozycji $k+1$ do w włącznie w obu elementach pary, tworząc ten sposób dwa nowe ciągi.

A	1	0	1	1	1
B	0	1	0	1	1
A'	1	0	0	1	1
B'	0	1	1	1	1

Krzyżowanie jednopunktowe

Krzyżowanie wielopunktowe jest uogólnieniem poprzednich operacji i charakteryzuje się odpowiednio większą liczbą l_k .

A	1	0	1	1	1
B	0	1	0	0	0
A'	1	1	0	1	1
B'	0	0	1	0	0

Krzyżowanie wielopunktowe

Krzyżowanie równomierne (jednolite, jednostajne) odbywa się zgodnie z wylosowanym wzorcem wskazującym, które geny dziedziczone są od pierwszego z rodziców (pozostałe od drugiego).

Krzyżowanie arytmetyczne. Dwuargumentowy operator jest zdefiniowany jako liniowa kombinacja dwóch wektorów. Jeżeli do krzyżowania zostały wybrane wektory (chromosomy) $x1$ i $x2$, to potomkowie są wyznaczani następująco:

$$x1' = a x1 + (1-a) x2$$

$$x2' = a x2 + (1-a) x1.$$

a jest wartością z przedziału $[0,1]$, co gwarantuje, że potomkami będą rozwiązania dopuszczalne.

Krzyżowanie heurystyczne. Ten operator do kierunku poszukiwań używa wartości funkcji celu, tworzy tylko jednego potomka i może w ogóle nie utworzyć potomka. Nowego osobnika tworzy się w następujący sposób:

$$ch3 = r(ch2 - ch1) + ch2,$$

gdzie r jest wartością losową z przedziału $[0,1]$ i $F(ch2) \leq F(ch1)$

Mutacja (mutation) jest zamianą pojedynczego genu na przeciwny z prawdopodobieństwem $0 \leq p_m \leq 0,1$. Wyróżnia się dwa rodzaje mutacji.

Pierwszego stopnia polega na tym, że każdy gen podlega losowaniu czy zostanie zmutowany.

Drugiego stopnia polega na tym, że losuje się czy dany osobnik zostanie zmutowany a następnie każdy gen podlega losowaniu, czy będzie mutowany.

Prawdopodobieństwo mutacji jest równa iloczynowi prawdopodobieństwa mutacji osobnika oraz prawdopodobieństwa mutacji genu.

A'	1	0	0	1	1
A''	1	0	1	1	1
Mutacja					

Nową populację tworzą chromosomy powstałe w wyniku selekcji i działania operatorów genetycznych. Nowa populacja zastępuje w całości starą, i staje się bieżącą w kolejnej iteracji algorytmu genetycznego.

Podstawowa forma mutacji może być zapisana następująco: $x' = m(x)$

gdzie x jest chromosomem rodzica, m funkcją mutacji, a x' chromosomem potomka[3]. Jeśli w przypadku chromosomów zakodowanych binarnie nie ma problemu ze stosowaniem mutacji (po prostu zamieniamy wartość genu na przeciwny), to w przypadku wartości zmiennopozycyjnych nie jest to już takie oczywiste. Poniżej przedstawiono kilka koncepcji na implementację mutacji.

Mutacja równomierna. Operator ten operując na jednym rodzicu x tworzy potomka x' . Operator wybiera losowo pozycję genu, który zostanie mutowany k ($1, \dots, q$). Mutowany wektor $x = (x_1, \dots, x_k, \dots, x_q)$ przekształca się w nowy $x' = (x_1, \dots, x_k', \dots, x_q)$, gdzie x_k' przyjmuje wartość losową z dozwolonego przedziału dla tej wartości. Operator ten odgrywa szczególnie ważną rolę we wczesnych fazach procesu ewolucyjnego.

Mutacja brzegowa. Działa na podobnych postawach, co mutacja równomierna, natomiast x_k' może z

jednakowym prawdopodobieństwem przyjąć wartości brzegów dozwolonego przedziału. Operator ten jest przydatny dla zadań optymalizacji, gdy szukane rozwiązanie leży blisko brzegu dopuszczalnej przestrzeni poszukiwań.

Niech $P(0)$ oznacza początkową populację osobników, natomiast $P(k)$ populację bieżącą.

Z populacji $P(k)$ wybieramy metodą selekcji, do puli rodzicielskiej $M(k)$, chromosomy o najlepszym przystosowaniu. Dalej łączymy osobniki w pary i dokonując operacji

krzyżowania zgodnie z prawdopodobieństwem krzyżowania p_c oraz mutacji zgodnie z prawdopodobieństwem p_m , otrzymujemy nową populację $P(k+1)$, do której wchodzi potomkowie populacji $M(k)$. Dla danego schematu chcemy, aby liczba chromosomów pasujących do danego schematu w populacji $P(k)$ wzrastała ze wzrostem liczby iteracji k .

IV Przykłady działania Algorytmów Genetycznych

Przykład 1

Rozważmy problem znalezienia maksimum funkcji $f(x)=2x^2$.

Do kodowania użyjemy pięciopozycyjnego systemu dwójkowego. Losowo wybierzemy początkową populację złożoną z czterech ciągów kodowych.

nr	populacja	wartość x	f(x)	pselect	oczekiwana liczba kopi	liczba wylosowanych	kopi
1	01101	13	338	0.14	0.58	1	
2	11000	24	1152	0.49	1.97	2	
3	01000	8	128	0.06	0.22	0	
4	10011	19	722	0.3	1.23	1	
suma			2340	1	4	4	
średnia			585	0.25	1	1	
maksimum			1152	0.49	1.97	2	

pula rodzicielska	partner	punkt k	nowa populacja	wartość x	f(x)
01101	2	4	01100	12	288
11000	1	4	11001	25	1250
11000	4	2	11011	27	1458
10011	3	2	10000	16	512
suma					3508
średnia					877
maksimum					1458

Przykład 2

Rozważmy teraz funkcję przystosowania, która określa liczbę jedynek w chromosomie. Chromosom ma 12 genów, a populacja liczy 8 osobników.

nr	populacja	wartość x	pselect	los. [0,100]	chrom osom	partne r	k	nowa populacja	f(x)
1	111001100101	7	0.152	79	7	3	5	101011110011	8
2	001100111010	6	0.130	44	3	4	1	011101110010	7
3	011101110011	8	0.174	9	7	1	3	101001100101	6
4	001000101000	3	0.065	74	1	7	3	111011011011	9
5	010001100100	4	0.087	44	3	7	4	011111011011	9
6	010011000101	5	0.108	86	7	2	5	101010111010	7
7	101011011011	8	0.173	48	4	3	1	001000101001	4
8	000010111100	5	0.108	23	2	7	4	001111011011	8
suma		46							58
średnia		6							7.25
maks		8							9

V Kodowanie

W klasycznym algorytmie stosuje się kodowanie binarne chromosomów.

Poszukujemy maksimum funkcji przystosowania $f(x_1, x_2, x_3, \dots, x_n) > 0$ dla $a < x_i < b$ a rozwiązanie ma mieć q miejsc po przecinku, wtedy przedział dzielimy na $(b-a) \cdot 10^q$ przedziałów, o $r=10^{-q}$, natomiast wymaganą długość ciągu wyraża wzór $(b_i - a_i) \cdot 10^q = \leq 2^{m_i} - 1$.
 $x_i = a_i + y_i (b_i - a_i) / (2^{m_i} - 1)$ jest wartość x_i , gdzie y_i jest wartością dziesiętną.

Można stosować także *kod Gray'a*, który charakteryzuje się tym, że ciągi binarne odpowiadające dwóm kolejnym liczbom całkowitym różnią się jednym bitem.

Innym rodzajem jest kodowanie logarytmiczne używane jest w celu zmniejszenia długości chromosomów w algorytmie genetycznym. Pierwszy bit (a) ciągu kodowanego jest bitem znaku funkcji wykładniczej, drugi bit (b) jest bitem znaku wykładnika funkcji wykładniczej, a pozostałe bity (bin) są reprezentacją wykładnika funkcji wykładniczej gdzie $[bin]_{10}$ oznacza wartość dziesiętną liczby zakodowanej w postaci ciągu binarnego bin.

$$[ab \text{ bin}] = (-1)^b e^{(-1)^a [bin]_{10}}$$

Przykłady

10110 jest równe jest liczbie $(-1)^0 e^{(-1)^1 [110]_{10}} = e^{(-6)} = 0.0024$

01010 jest równe jest liczbie $(-1)^1 e^{(-1)^0 [010]_{10}} = -e^{-2} = -0.0024$

VI Schematy

Schemat jest zbiór chromosomów o pewnych wspólnych cechach. Do opisywania schematów używa się metasymbolu *. Przykładowo schemat *00*1 opisuje następujące ciągi {00001, 00011, 10001, 10011}.

Wyznaczenie wszystkich możliwych schematów k -elementowego alfabetu określa się wzorem: $(k+1)^w$, gdzie w jest długością słowa.

Mówimy że chromosom należy do danego schematu (jest reprezentantem schematu), jeżeli dla każdej pozycji (locus) $j = 1, 2, \dots, l$, gdzie l jest długością chromosomu, symbol występujący na j -tej pozycji chromosomu odpowiada symbolowi na j -tej pozycji schematu.

Rzędem schematu nazywam liczbę ustalonych pozycji we wzorcu. Rząd schematu będzie oznaczać przez $o(H)$

Rozpiętością schematu nazywam odległość między dwoma skrajnymi pozycjami ustalonymi. Rozpiętość schematu będzie oznaczać przez $d(S)$.

schemat	rząd	rozpiętość
****0	1	0
1***1	2	4
000*1	4	4
**1*0	2	2

Badanie rzędu i rozpiętości schematu

Podobnie jak w przypadku chromosomów, tak w przypadku schematów mówi się o ich przystosowaniu. Przystosowanie schematu jest średnim przystosowaniem obliczonym na podstawie przystosowania wszystkich jego reprezentantów.

genotyp	ocena	schemat	reprezentant	ocena
11010	26	a 1***0	a, c, f	$(26+28+16)/3=23.3$

10111	23	b	*0**1	b, d, e	10.3
11100	28	c	**1**	b, c, d	25.3
00101	5	d	00***	d, e	4
00011	3	e			
10000	16	f			

Ocena schematu

Średnie przystosowanie osobników w tej populacji wynosi 16,8. Najlepiej przystosowany jest osobnik z oceną 3 (minimalizujemy funkcję oceny), a najgorzej osobnik z oceną 28.

Podobnie jak w ciągach za przetwarzanie schematów w algorytmie genetycznym mają wpływ: selekcja, krzyżowanie i mutacja.

Niech S będzie danym schematem, $c(S, k)$ ilością chromosomów w populacji $P(k)$ pasujących do schematu S . Zatem $c(S, k)$ jest liczbą elementów zbioru $P(k) \cap S$.

$F(S, k)$ oznacza średnią wartość funkcji przystosowania chromosomów w populacji $P(k)$, pasujących do schematu S . $F(S, k)$ nazywa się również przystosowaniem schematu S w iteracji k .

$\square(k)$ oznacza sumę wartości funkcji przystosowania chromosomów w populacji $P(k)$ o liczebności N .

F_{sr} średnią wartość funkcji przystosowania chromosomów w danej populacji.

$chr^{(k)}$ będzie elementem puli rodzicielskiej $M(k)$.

Dla każdego $chr^{(k)} \in M(k)$ i dla każdego $i = 1, \dots, c(S, k)$ prawdopodobieństwo, że $chr^{(k)} = ch_i$ dane jest wzorem: $F(ch_i) / \square(k)$

Skoro każdy chromosom z puli rodzicielskiej $M(k)$ jest równocześnie chromosomem należącym do populacji $P(k)$, to chromosomy ze zbioru $M(k) \cap S$ są po prostu tymi samymi chromosomami, które ze zbioru $P(k) \cap S$ zostały wybrane do populacji $M(k)$. Oznaczając przez $b(S, k)$ ilość chromosomów z puli rodzicielskiej $M(k)$ pasujących do schematu S z powyższych rozważań otrzymujemy następujące wnioski

1. Jeżeli schemat S zawiera chromosomy o wartości funkcji przystosowania powyżej średniej, czyli przystosowanie schematu S w iteracji k jest większe niż średnia wartość funkcji przystosowania chromosomów w populacji $P(k)$, to oczekiwana liczba chromosomów pasujących do schematów puli rodzicielskiej $M(k)$ jest większa niż ilość chromosomów pasujących do schematu S w populacji $P(k)$.
2. Dla danego chromosomu w $M(k) \cap S$ prawdopodobieństwo, że chromosom ten zostanie wybrany do krzyżowania i żaden z jego potomków nie będzie należał do schematu S jest ograniczone z góry przez *prawdopodobieństwo zniszczenia schematu S*
3. Dla danego chromosomu w $M(k) \cap S$ prawdopodobieństwo, że chromosom ten nie zostanie wybrany do krzyżowania albo, co najmniej jeden z jego potomków będzie należał do schematu S po krzyżowaniu jest ograniczone z dołu przez *prawdopodobieństwo przetrwania schematu S* .
4. Dla danego chromosomu w $M(k) \cap S$ prawdopodobieństwo, że chromosom ten będzie należał do schematu S po operacji mutacji, jest dane przez $(1 - p_m)^{o(S)}$
Wielkość tę nazywamy *prawdopodobieństwem przetrwania mutacji* przez schemat S .

5. Jeżeli prawdopodobieństwo mutacji p_m jest małe, to można przyjąć, że prawdopodobieństwo przetrwania mutacji przez schemat S , określone w poprzednim wniosku jest w przybliżeniu równe $1 - p_m o(S)$

Twierdzenie o schematach. Schematy małego rzędu, o małej rozpiętości i o przystosowaniu powyżej średniej otrzymują rosnącą wykładniczo liczbę swoich reprezentantów w kolejnych generacjach algorytmu genetycznego.

VII Zastosowania algorytmów genetycznych

Przykładem zastosowania algorytmów genetycznych jest problem komiwojażera, istota problemu polega na znalezieniu najkrótszej drogi łączącej wszystkie miasta, tak by przez każde miasto przejść tylko raz. Algorytmy genetyczne równie dobrze radzą sobie w znajdowaniu przybliżeń ekstremów funkcji, których nie da się obliczyć analitycznie.

Wykorzystywane są także do zarządzania populacją sieci neuronowych, projektowania maszyn. Na algorytmach genetycznych bazuje się przy projektowaniu obwodów elektrycznych.

Główna różnica tkwi w algorytmie budowy osobnika na podstawie genomu. Ma on postać instrukcji dla programu, który na jego podstawie buduje obwód elektryczny. Najpierw mamy proste połączenie wejścia z wyjściem. Następnie program dodaje i usuwa połączenia i elementy. Zbudowany tak obwód jest oceniany na podstawie prostych zależności fizycznych.

Algorytmy genetyczne używa się w połączeniu z układami FPGA (field-programmable gate arrays)(układ scalony, którego działanie może być określone przez użytkownika przy użyciu programatora). Algorytmy genetyczne badają zwykle zachowanie symulowanych pokoleń. Dzięki układom FPGA możliwe jest ewoluowanie prawdziwych obwodów elektrycznych. Są one wpisywane do chipa, a następnie ich właściwości elektryczne są mierzone rzeczywistym obwodem testowym. W ten sposób ewolucja może wykorzystać wszystkie fizyczne właściwości rzeczywistego układu elektrycznego.

Regulatory stosowane w automatyce udoskonala się o algorytmy genetyczne. Najpopularniejszy algorytm sterowania, czyli PID jest rodzajem zestawu połączonych ze sobą członów różniczkujących i całkujących. Odpowiedni algorytm genetyczny może zbudować taki układ analogicznie do obwodu elektrycznego. Korzystając z tej metody John R. Koza opracował nowe wersje PID-a (<http://www.genetic-programming.com/hc/pid.html>)

Bibliografia

1. Goldberg D.E.: Algorytmy genetyczne i ich zastosowanie, WNT, Warszawa.1995.
2. Rutkowska D., Piliński M., Rutkowski L.: Sieci neuronowe, algorytmy genetyczne i systemy rozmyte, PWN, Warszawa, 1999
3. Łacki W.: Wykorzystanie inteligentnych technik obliczeniowych w zarządzaniu projektem
4. Chodak G., Kwaśnicki W., Zastosowanie algorytmów genetycznych w prognozowaniu popytu
5. <http://panda.bg.univ.gda.pl/%7Esielim/genetic/index.htm>