

Podstawowy algorytm genetyczny

Dr hab. inż. Katarzyna Pentoś

Algorytmy ewolucyjne – definicje

Populacja - zbiór osobników o określonej liczbeności.

Osobnik - zakodowany w postaci chromosomów zbiór parametrów zadania, czyli rozwiązanie, określone też jako punkt przestrzeni poszukiwań.

Chromosom – łańcuch lub ciąg kodowy – uporządkowany ciąg genów.

Gen – pojedynczy element genotypu, w szczególności chromosomu.

Genotyp - zespół chromosomów danego osobnika. Osobnikami populacji mogą być genotypy albo pojedyncze chromosomy (jeśli genotyp składa się tylko z jednego chromosomu).

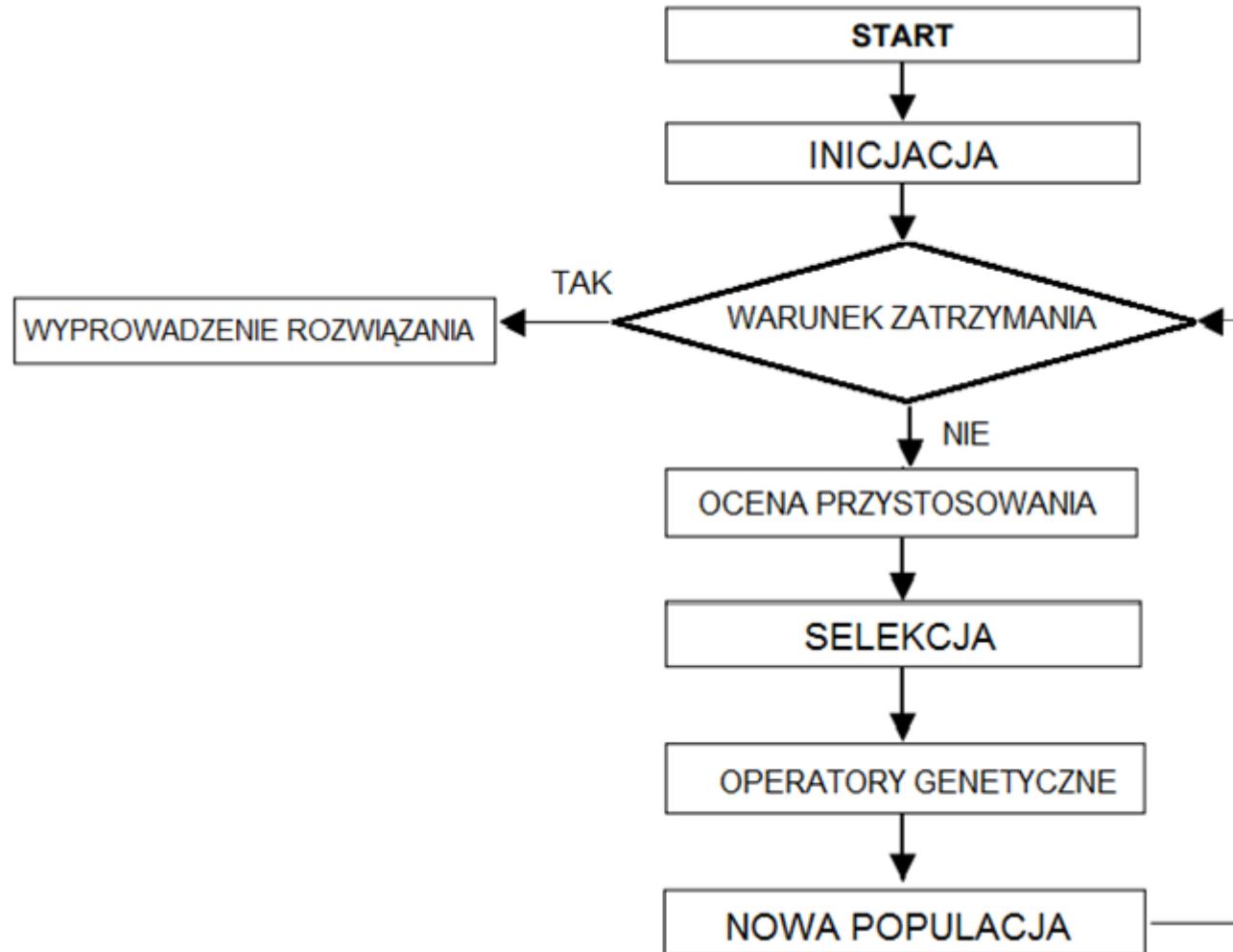
Fenotyp - zestaw wartości odpowiadających danemu genotypowi, zdekodowana struktura będąca wprost rozwiązaniem zadania/problemu.

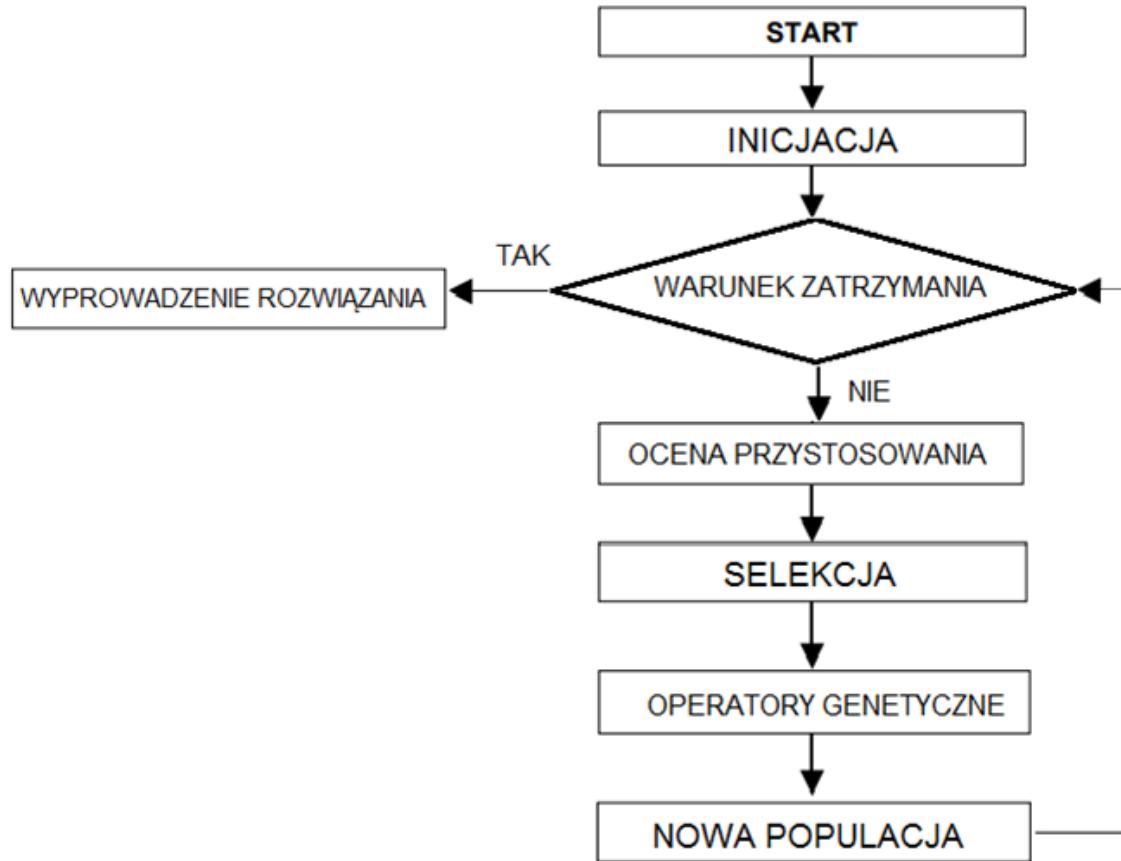
Allel - wartość danego genu, określona jako wartość cechy lub wariant cechy.

Locus (pozycja) - wskazuje miejsce położenia danego genu w łańcuchu, czyli chromosomie.

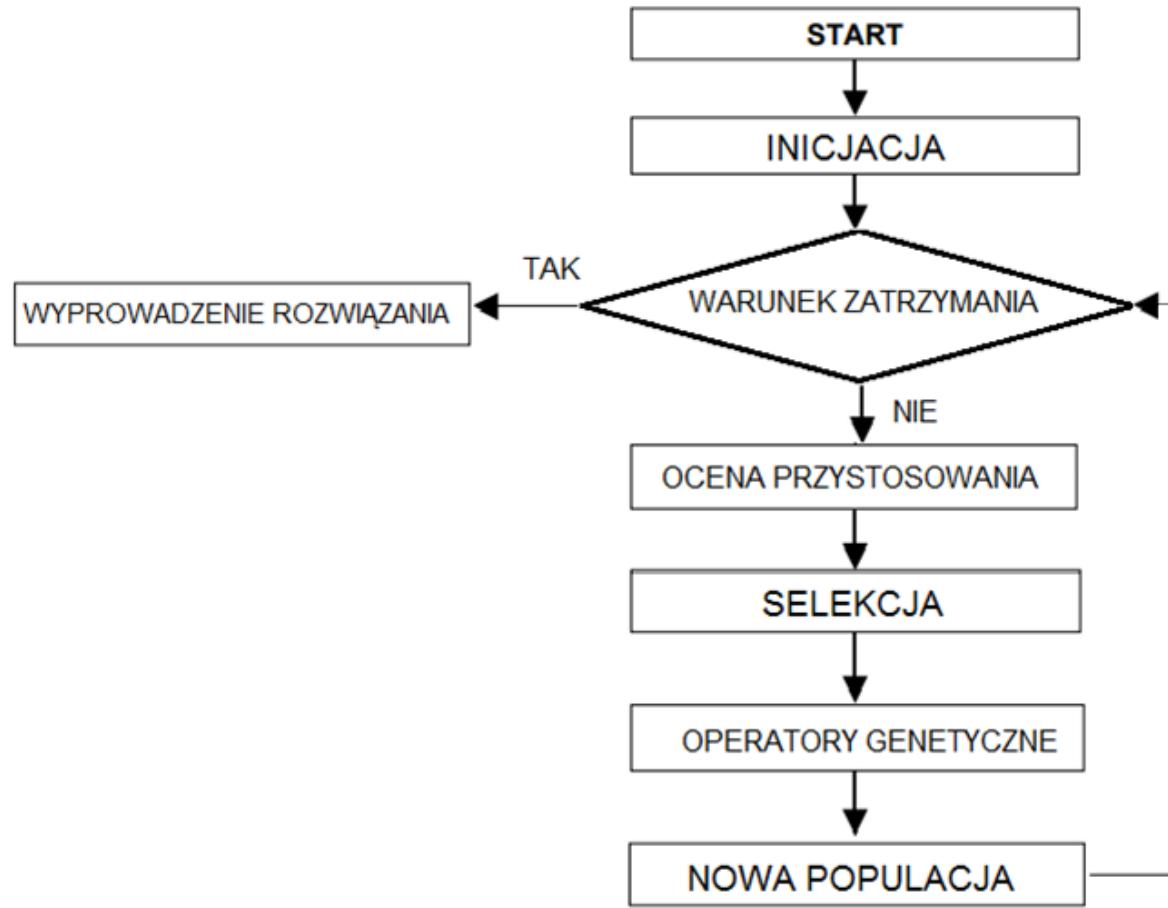
Funkcja przystosowania (funkcja dopasowania, funkcja oceny) - stanowi ona miarę przystosowania (dopasowania) danego osobnika w populacji.

Klasyczny algorytm genetyczny

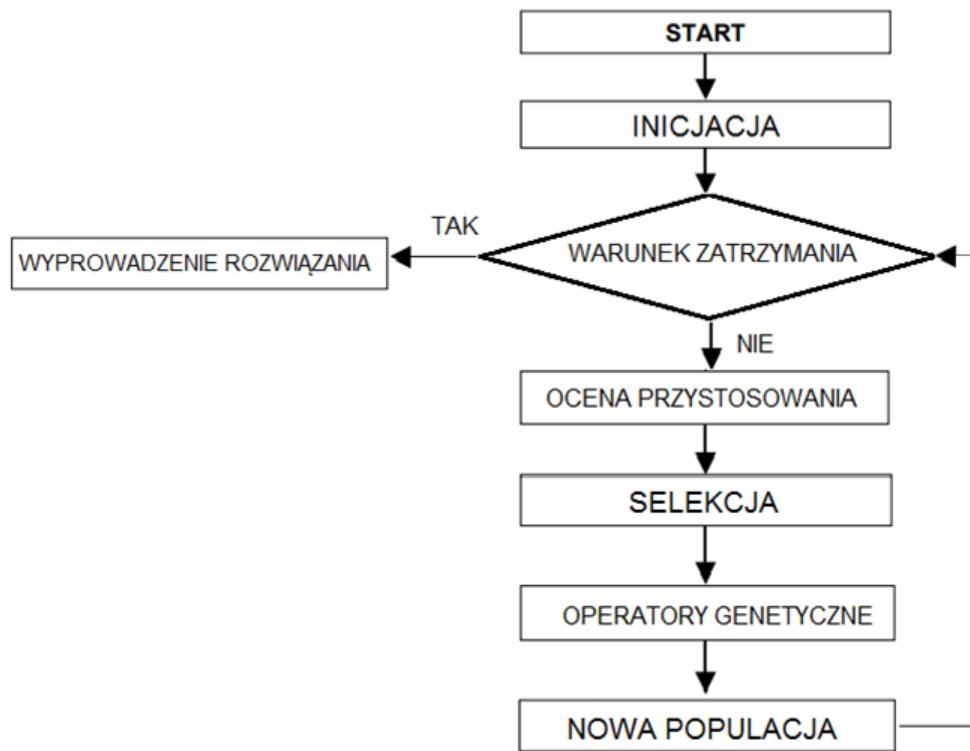




1. Inicjacja, czyli utworzenie populacji początkowej, polega na losowym wyborze żądanej liczby chromosomów (osobników) reprezentowanych przez ciągi binarne o określonej długości.

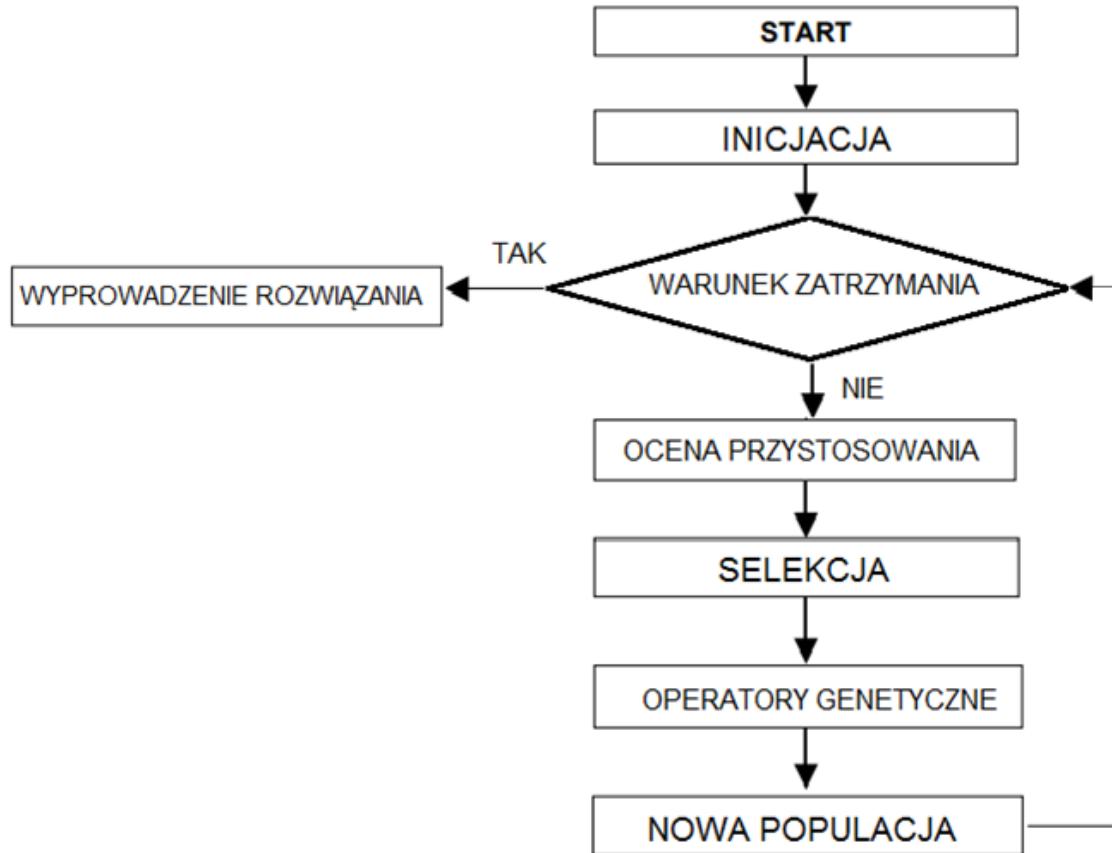


2. Ocena przystosowania chromosomów – obliczenie wartości funkcji przystosowania dla każdego z chromosomów. Postać funkcji zależy od rozwiązywanego problemu, przyjmuje zawsze wartości nieujemne.

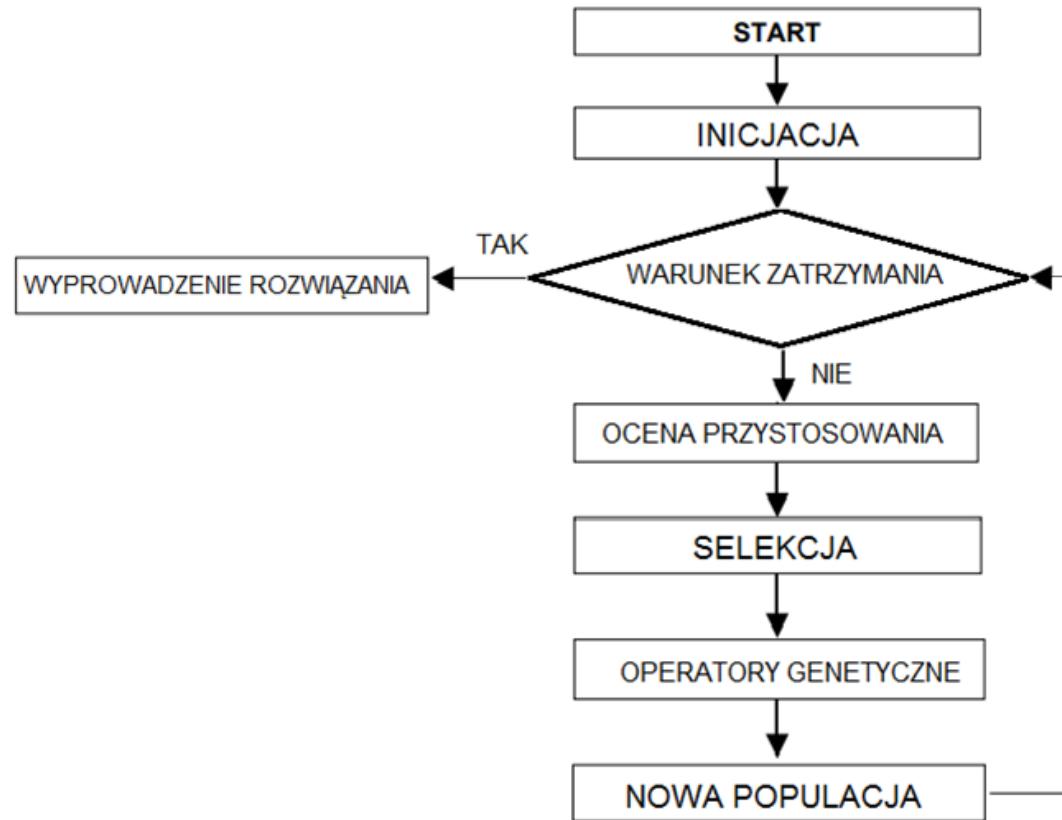


3. Selekcja chromosomów polega na wybraniu (na podstawie wartości funkcji przystosowania), tych chromosomów, które będą brały udział w tworzeniu potomków do następnej generacji.

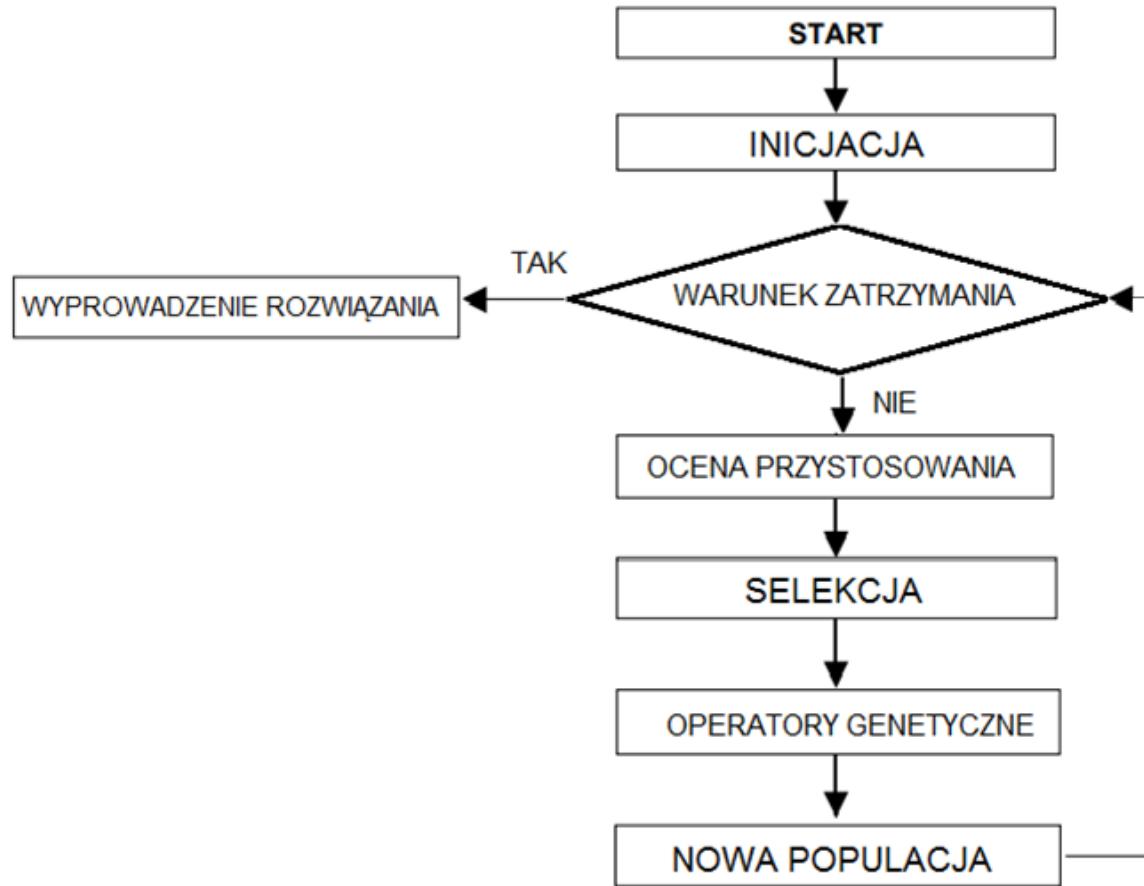
W wyniku procesu selekcji powstaje populacja rodzicielska **o takiej samej liczbeności jak bieżąca populacja**.



4. Utworzenie nowej populacji rodzicielskiej po zastosowaniu operatorów krzyżowania i mutacji.



5. Sprawdzenie warunku zatrzymania. Warunek zatrzymania to może być określona wartość błędu, sytuacja gdy dalsze działanie algorytmu nie poprawia uzyskanej, najlepszej wartości, przekroczenie określonego czasu działania lub liczby iteracji algorytmu.



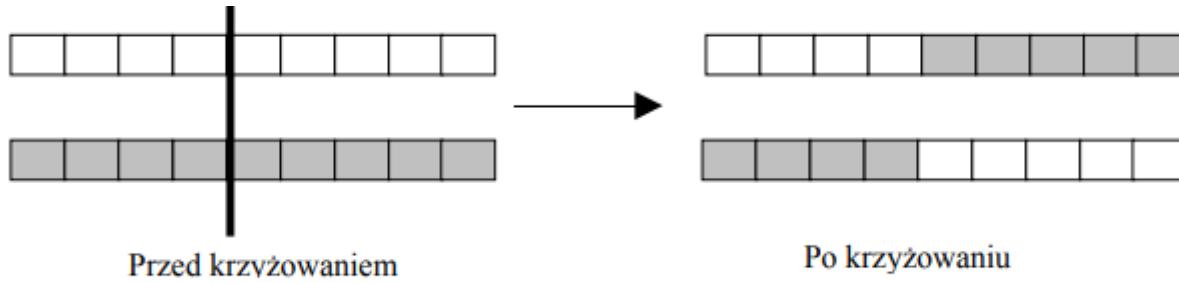
6. Wyprowadzenie najlepszego chromosomu. Po spełnieniu warunku zatrzymania należy wyprowadzić najlepszy chromosom czyli podać rozwiązanie problemu. Najlepszym rozwiązaniem jest chromosom o największej wartości funkcji przystosowania.

Operatory genetyczne

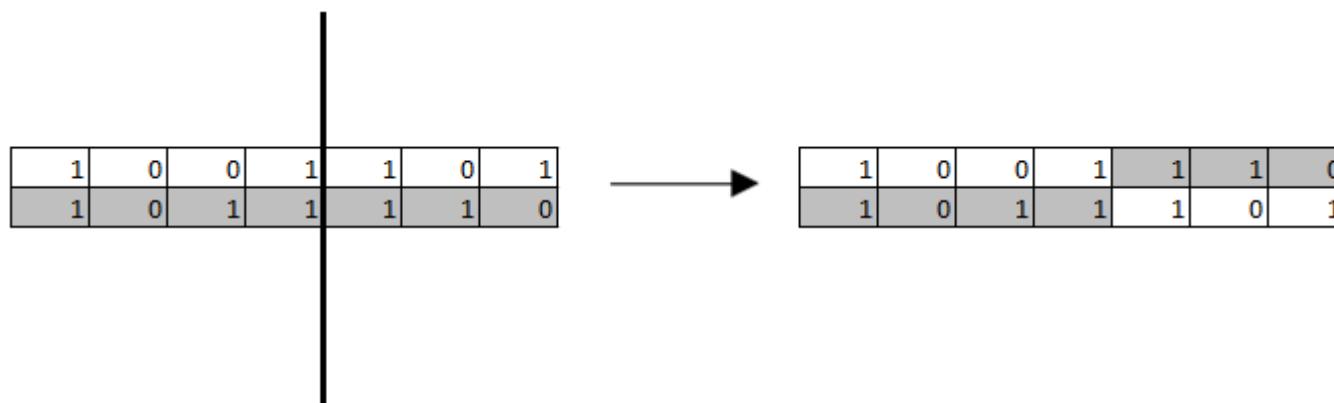
Krzyżowanie

- Wybieramy pulę rodzicielską i kojarzymy chromosomy w pary
- Losujemy pozycję genu (locus) w chromosomie określającą punkt krzyżowania. Jeśli każdy z rodziców składa się z L genów to punkt krzyżowania jest liczbą I z zakresu $[1, L-1]$.
- W wyniku krzyżowania otrzymuje się parę potomków:
 1. Potomek, którego chromosom składa się z genów na pozycjach od 1 do I pochodzących od pierwszego rodzica i pozostałych genów pochodzących od drugiego rodzica
 2. Potomek, którego chromosom składa się z genów na pozycjach od 1 do I pochodzących od drugiego rodzica i pozostałych genów pochodzących od pierwszego rodzica

Krzyżowanie jednopunktowe



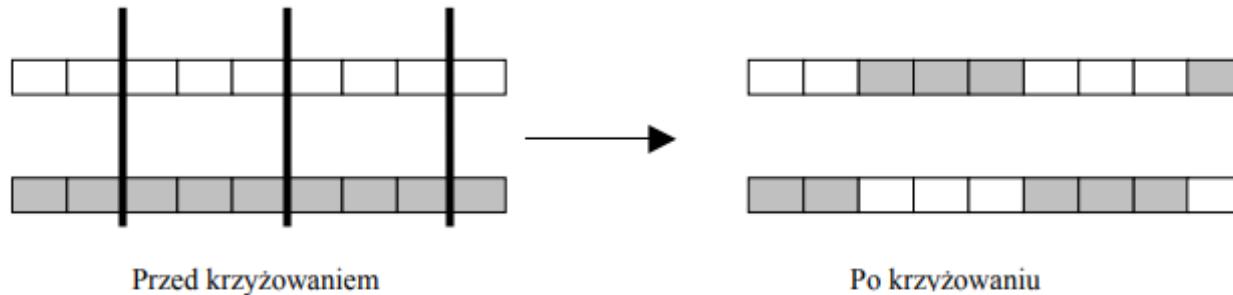
Przykład dla $|l|=4$



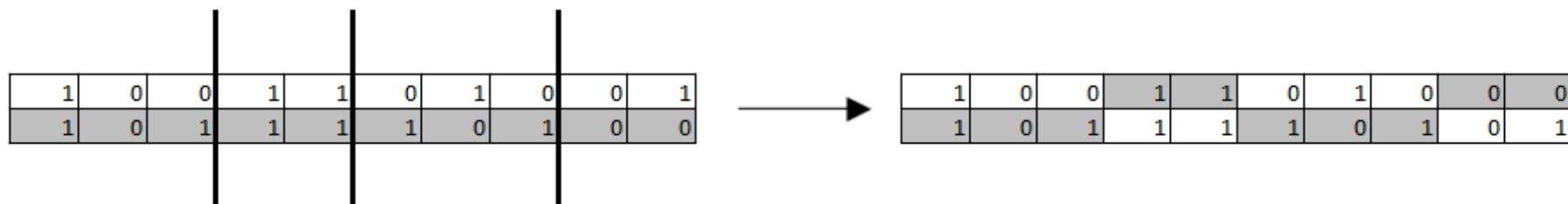
Krzyżowanie wielopunktowe

Wybierane są dwa lub więcej punkty krzyżowania chromosomów

Usprawnia proces krzyżowania w przypadku korzystania z długich chromosomów



Przykład dla $l=3$, $l=5$ i $l=8$



Mutacja

Zmienia wartość wybranego losowo genu w chromosomie na przeciwny (1 na 0, 0 na 1)

Mutacja zachodzi bardzo rzadko – prawdopodobieństwo mutacji jest zwykle dużo mniejsze niż krzyżowania

Celem mutacji jest wprowadzenie różnorodności populacji

Przykład dla l=5



Metody selekcji

Metoda koła ruletki

Selekcja dokonuje się, poprzez wybór chromosomów, którym na kole (koło ruletki) przydzielono sektory proporcjonalne do wartości przystosowania

Większa wartość przystosowania = częstszy wybór do populacji rodzicielskiej

Lepiej przystosowane chromosomy mogą być wybierane wielokrotnie

Skutek: miejsce „słabszych” zajmują „mocniejsi”

Wielkość sektorów na *kole ruletki* przydzielane są następująco:

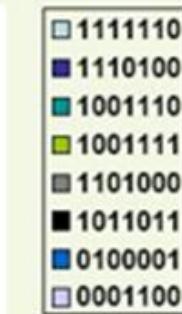
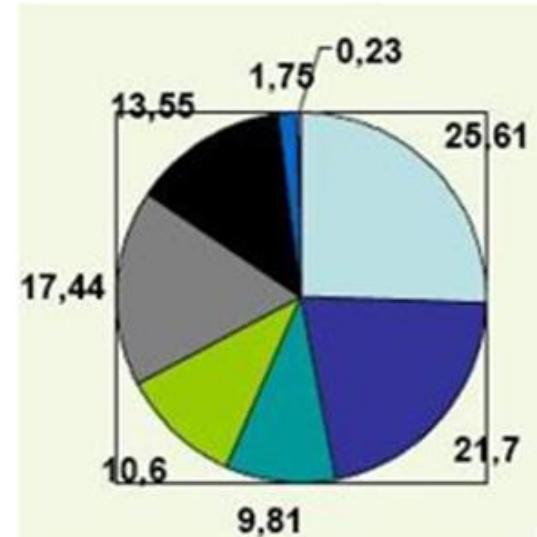
$$p(ch_i) = \frac{F(ch_i)}{\sum_{i=1}^n F(ch_i)} \cdot 100\%$$

Oznacza to iloraz funkcji przystosowania danego osobnika i sumy funkcji przystosowania osobników w całej populacji wyrażony w %

Metoda koła ruletki – przykład

chromosom	fenotyp	funkcja przystosowania	wielkość procentowa sektora
1111110	(126)	31754	25,61%
1110100	(116)	26914	21,70%
1001110	(78)	12170	9,81%
1001111	(79)	12484	10,6%
1101000	(104)	21634	17,44%
1011011	(91)	16564	13,35%
0100001	(33)	2180	1,75%
0001100	(12)	290	0,23%

Suma: 123 990



Algorytm genetyczny – przykład 1

Szukanie maksimum funkcji $y=2x+1$ dla $x \in [0,31]$

x – parametr zadania.

Zbiór $\{0,1,2,\dots,31\}$ – przestrzeń poszukiwań a jednocześnie zbiór potencjalnych rozwiązań zadania

Rozwiązania kodujemy binarnie za pomocą 5 bitów.

Ciągi kodowe to chromosomy a w tym przypadku także genotypy.

Jako funkcję przystosowania przyjmiemy **y=2x+1**

1. Losujemy populację początkową

W wyniku losowania otrzymujemy:

$$ch_1 = [00110]$$

$$ch_2 = [00101]$$

$$ch_3 = [01101]$$

$$ch_4 = [10101]$$

$$ch_5 = [11010]$$

$$ch_6 = [10010]$$

$$ch_7 = [01000]$$

$$ch_8 = [00101]$$

co odpowiada fenotypom:

$$ch_1^* = 6$$

$$ch_2^* = 5$$

$$ch_3^* = 13$$

$$ch_4^* = 21$$

$$ch_5^* = 26$$

$$ch_6^* = 18$$

$$ch_7^* = 8$$

$$ch_8^* = 5$$

2. Obliczamy funkcję przystosowania

$$F(ch_1) = 2 \cdot ch_1^* + 1 = 13$$

$$F(ch_2) = 11$$

$$F(ch_3) = 27$$

$$F(ch_4) = 43$$

$$F(ch_5) = 53$$

$$F(ch_6) = 37$$

$$F(ch_7) = 17$$

$$F(ch_8) = 11$$

$$\text{Suma} = 212$$

3. Selekcja chromosomów – koło ruletki

Na podstawie wzoru:

$$p(ch_i) = \frac{F(ch_i)}{\sum_{i=1}^n F(ch_i)} \cdot 100\%$$

obliczamy wycinki koła ruletki wyrażone w procentach:

$$v(ch_1) = (13/212) \cdot 100\% = 6,13\%$$

$$v(ch_2) = 5,19\%$$

$$v(ch_3) = 12,74\%$$

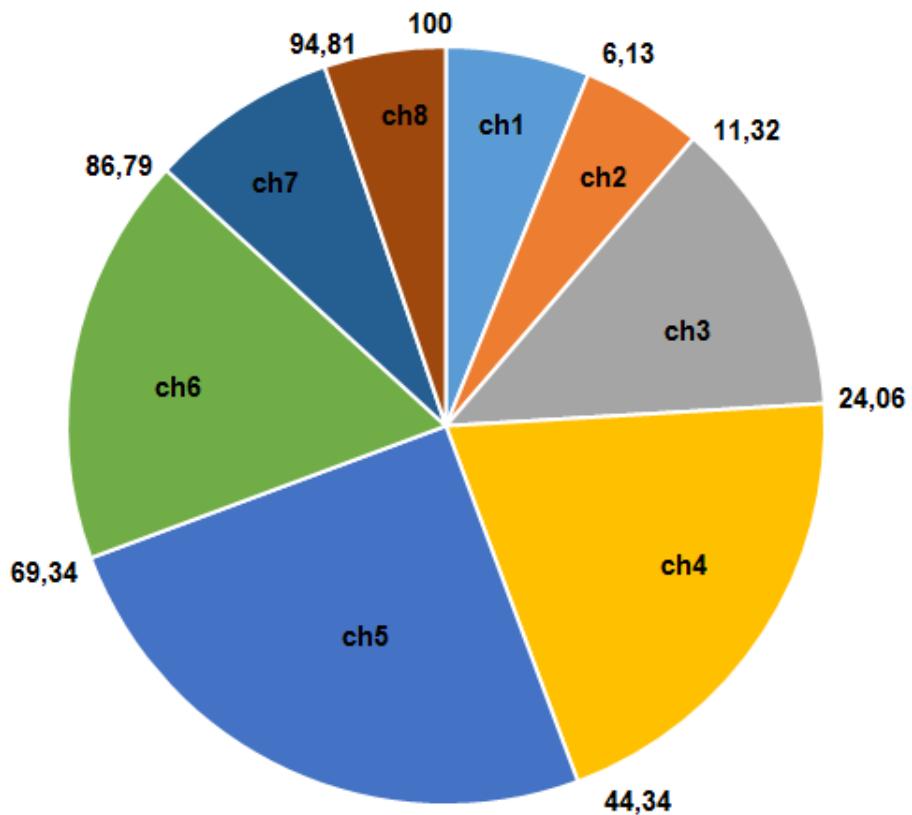
$$v(ch_4) = 20,28\%$$

$$v(ch_5) = 25\%$$

$$v(ch_6) = 17,45\%$$

$$v(ch_7) = 8,02\%$$

$$v(ch_8) = 5,19\%$$



3. Selekcja chromosomów – koło ruletki

Za pomocą koła ruletki losujemy 8 nowych chromosomów.

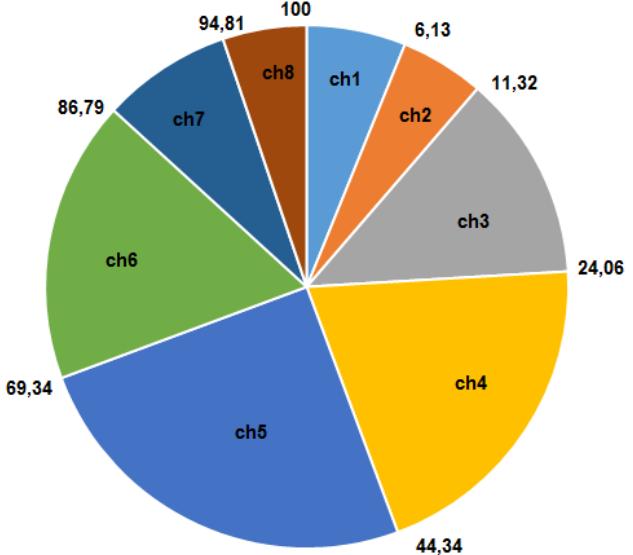
Założymy, że wylosowano następujące liczby:

79 44 9 74 45 86 48 23

Oznacza to wybór następujących chromosomów:

ch₆ ch₄ ch₂ ch₆ ch₅ ch₆ ch₅ ch₃

Te chromosomy tworzą pulę rodzicielską.



4. Operacje genetyczne

Załóżmy, że ustalono globalne prawdopodobieństwo mutacji $p_m=0,2$ i prawdopodobieństwo krzyżowania $p_k=0,75$

Krzyżowanie:

Kojarzymy osobniki w pary tak jak są ułożone w puli rodzicielskiej. Dla każdej pary losujemy liczbę z przedziału $[0,1]$. Jeśli dana liczba będzie mniejsza od $p_k=0,75$ to nastąpi krzyżowanie.

Załóżmy, że wylosowano: 0,12 0,73 0,65 0,95.

Oznacza to, że trzy pierwsze pary zostaną poddana krzyżowaniu a czwarta para nie.

Dodatkowo dla każdej pary podlegającej krzyżowaniu losujemy punkt krzyżowania (liczba całkowita z przedziału $[1,4]$).

Uzyskane wyniki:

Pierwsza para rodziców ($I_k=3$) Pierwsza para potomków

$ch_6=[100\mathbf{1}0]$ [1000 $\mathbf{1}$]

$ch_4=[1010\mathbf{1}]$ [10110]

Druga para rodziców ($I_k=4$) Druga para potomków

$ch_2=[0010\mathbf{1}]$ [00100]

$ch_6=[1001\mathbf{0}]$ [10011]

Trzecia para rodziców ($I_k=3$) Trzecia para potomków

$ch_5=[110\mathbf{1}0]$ [11010]

$ch_6=[100\mathbf{1}0]$ [10010]

Czwarta para rodziców Czwarta para potomków

$ch_5=[11010]$ [11010]

$ch_3=[0110\mathbf{1}]$ [01101]

Po operacji krzyżowania otrzymujemy następującą populację potomków o fenotypach

$$ch_1=[10001] \quad ch_1^*=17$$

$$ch_2=[10110] \quad ch_2^*=22$$

$$ch_3=[00100] \quad ch_3^*=4$$

$$ch_4=[10011] \quad ch_4^*=19$$

$$ch_5=[11010] \quad ch_5^*=26$$

$$ch_6=[10010] \quad ch_6^*=18$$

$$ch_7=[11010] \quad ch_7^*=26$$

$$ch_8=[01101] \quad ch_8^*=13$$

Mutacja

Dla każdego osobnika po krzyżowaniu losujemy liczbę z zakresu od 0 do 1. Założymy, że wylosowano

$ch_1=[10001]$ 0,56

$ch_2=[10110]$ 0,15

$ch_3=[00100]$ 0,48

$ch_4=[10011]$ 0,59

$ch_5=[11010]$ 0,06

$ch_6=[10010]$ 0,89

$ch_7=[11101]$ 0,39

$ch_8=[01010]$ 0,76

Mutacji podlegają te osobniki, dla których wylosowana liczba jest mniejsza niż prawdopodobieństwo mutacji $p_m=0,2$.

Dla osobników podlegających mutacji losujemy miejsce mutacji, liczbę całkowitą z zakresu [1,5]

ch₁=[10001] 0,56 NIE

ch₂=[10110] 0,15 TAK I=3

ch₃=[00100] 0,48 NIE

ch₄=[10011] 0,59 NIE

ch₅=[11010] 0,06 TAK I=2

ch₆=[10010] 0,89 NIE

ch₇=[11101] 0,39 NIE

ch₈=[01010] 0,76 NIE

Po operacji mutacji otrzymujemy następującą populację potomków o fenotypach

$$ch_1 = [10001] \quad ch_1^* = 17$$

$$ch_2 = [10010] \quad ch_2^* = 18$$

$$ch_3 = [00100] \quad ch_3^* = 4$$

$$ch_4 = [10011] \quad ch_4^* = 19$$

$$ch_5 = [10010] \quad ch_5^* = 18$$

$$ch_6 = [10010] \quad ch_6^* = 18$$

$$ch_7 = [11010] \quad ch_7^* = 26$$

$$ch_8 = [01101] \quad ch_8^* = 13$$

5. Obliczamy funkcje przystosowania dla nowej populacji

$$F(ch_1) = 2 \cdot ch_1^* + 1 = 35$$

$$F(ch_2) = 37$$

$$F(ch_3) = 9$$

$$F(ch_4) = 39$$

$$F(ch_5) = 37$$

$$F(ch_6) = 37$$

$$F(ch_7) = 59$$

$$F(ch_8) = 27$$

$$\text{Suma} = 280$$