Algorytmy Genetyczne (AG)

dr inż. Tomasz Białaszewski

Biologiczne tło AG

- **Genetyka** - badanie dziedziczności i zmienności organizmów
- Początek badań - XIX w. (Gregor Johann Mendel)
- **Gen** (z gr. *genos* - ród) - abstrakcyjna jednostki dziedziczności
- **Allel** – różne odmiany genu

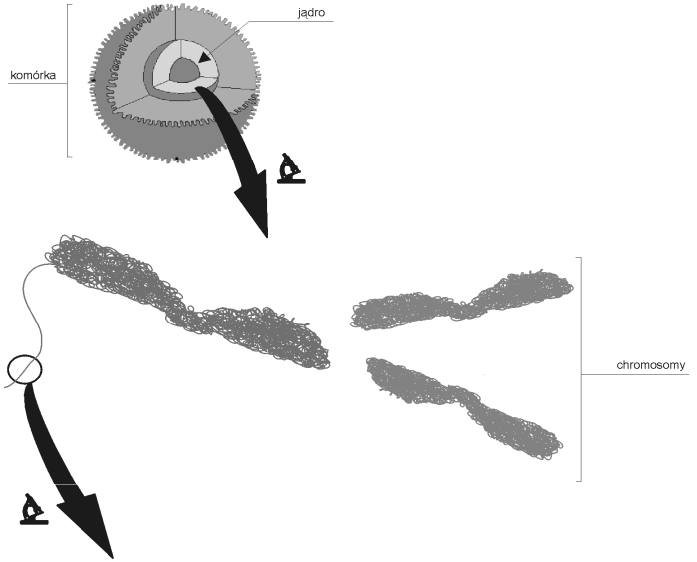
Biologiczne tło AG

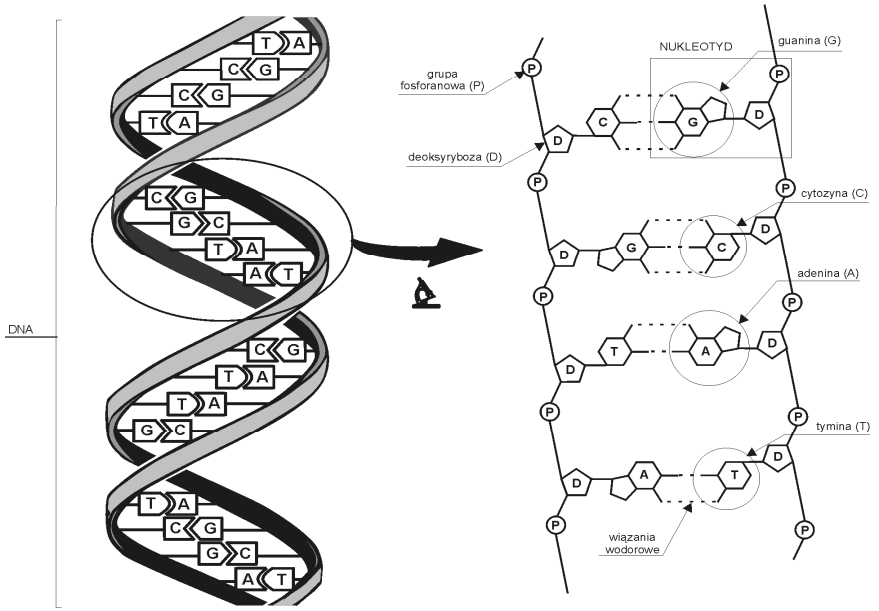
- **Chromosom** (z gr. *chromo* – barwa, *soma* – ciało) – struktury wewnątrz jąder komórek roślin i zwierząt zawierające geny
- Wszystkie komórki (z wyjątkiem rozrodczych) danego gatunku posiadają tę samą liczbę chromosomów
- **Diplont** – organizm posiadający pary homologiczne chromosomów
- U organizmów płciowych dwa niesparowane chromosomy nazywane (X i Y) określają płeć

Biologiczne tło AG

- Chromosom – **kwas dezoksyrybonukleinowy**

  (**DNA** – Friedrich Miescher, 1869).
- Podstawowe elementy DNA (nukleotydy):
  - Adenina (**A**)
  - Cytozyna (**C**)
  - Guanina (**G**)
  - Tymina (**T**)
- DNA - struktura podwójnej helisy (taśmy) (James Watson i Francis Crick 1953) – *komplementarne nici*



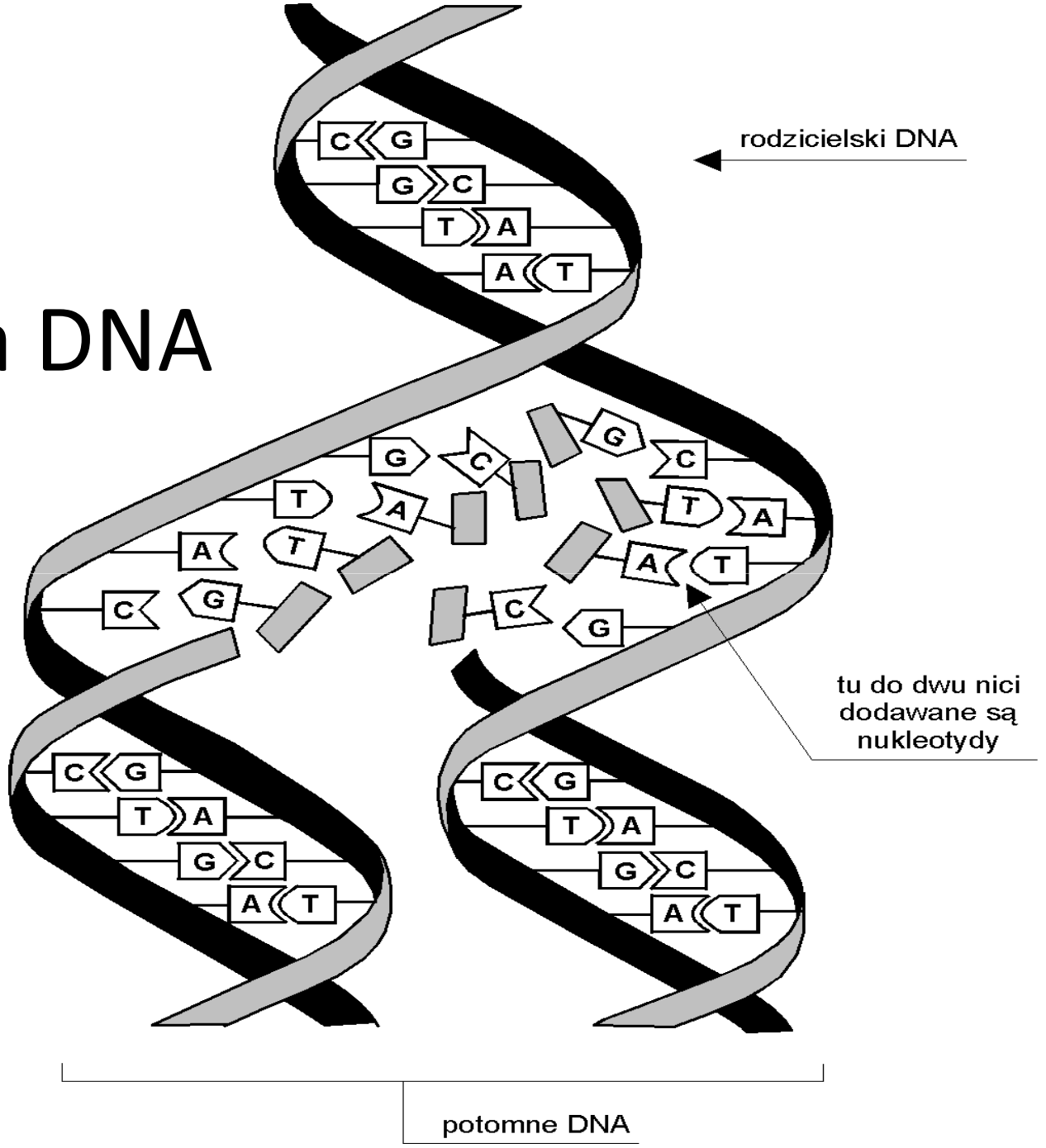


Biologiczne tło AG

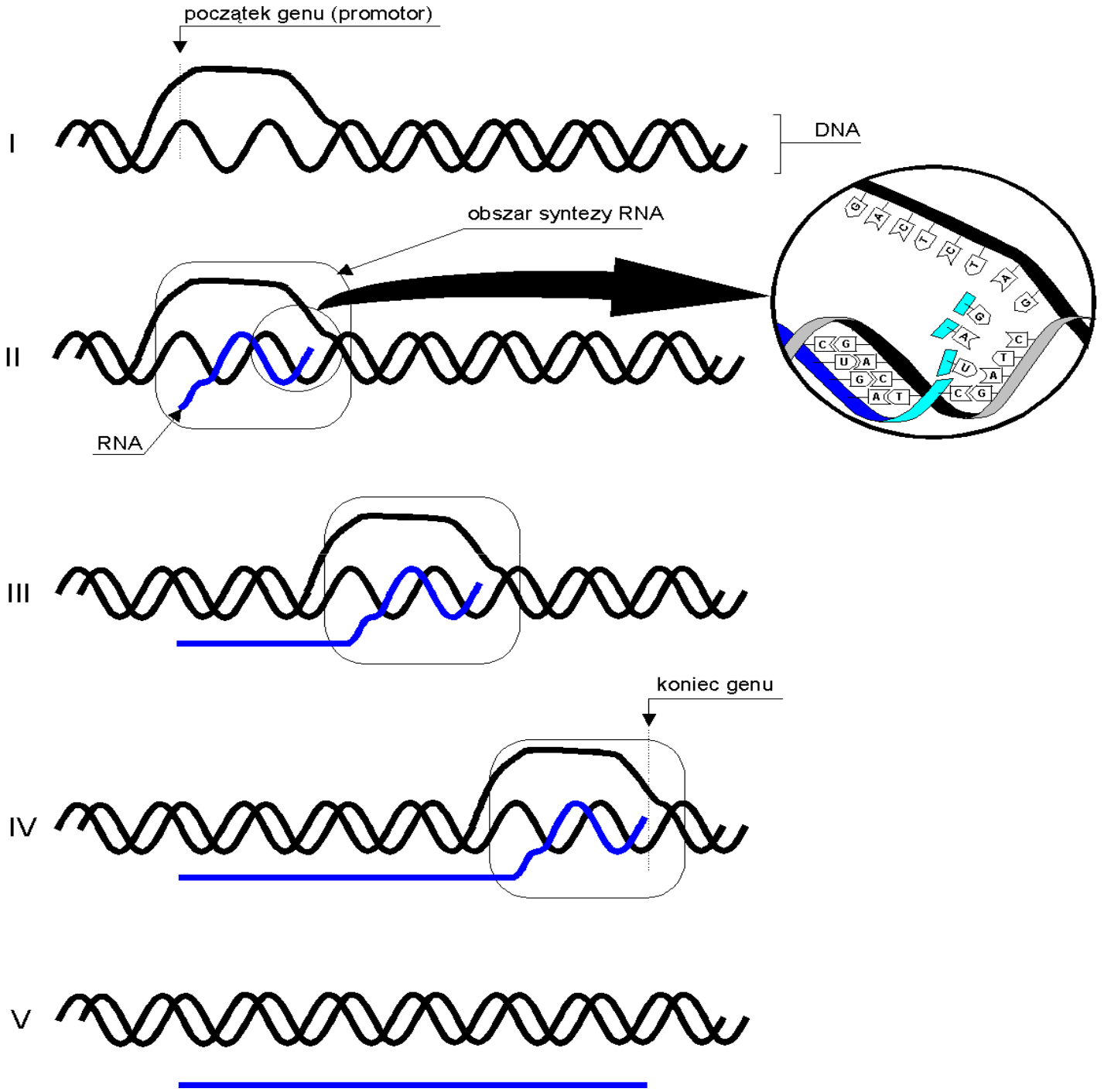
- **Replikacja** - przenoszenie i kopiowanie genów (DNA)

- Synteza informacji w genach:
  - **Transkrypcja** – przepisanie genów do ***kwasu rybonukleinowego*** (***RNA***)
  - **Translacja** - tłumaczenie fragmentów RNA na odpowiednie białka

Replikacja DNA



Synteza RNA - transkrypcja



Aminokwasowa interpretacja kodonów nici informacyjnej mRNA

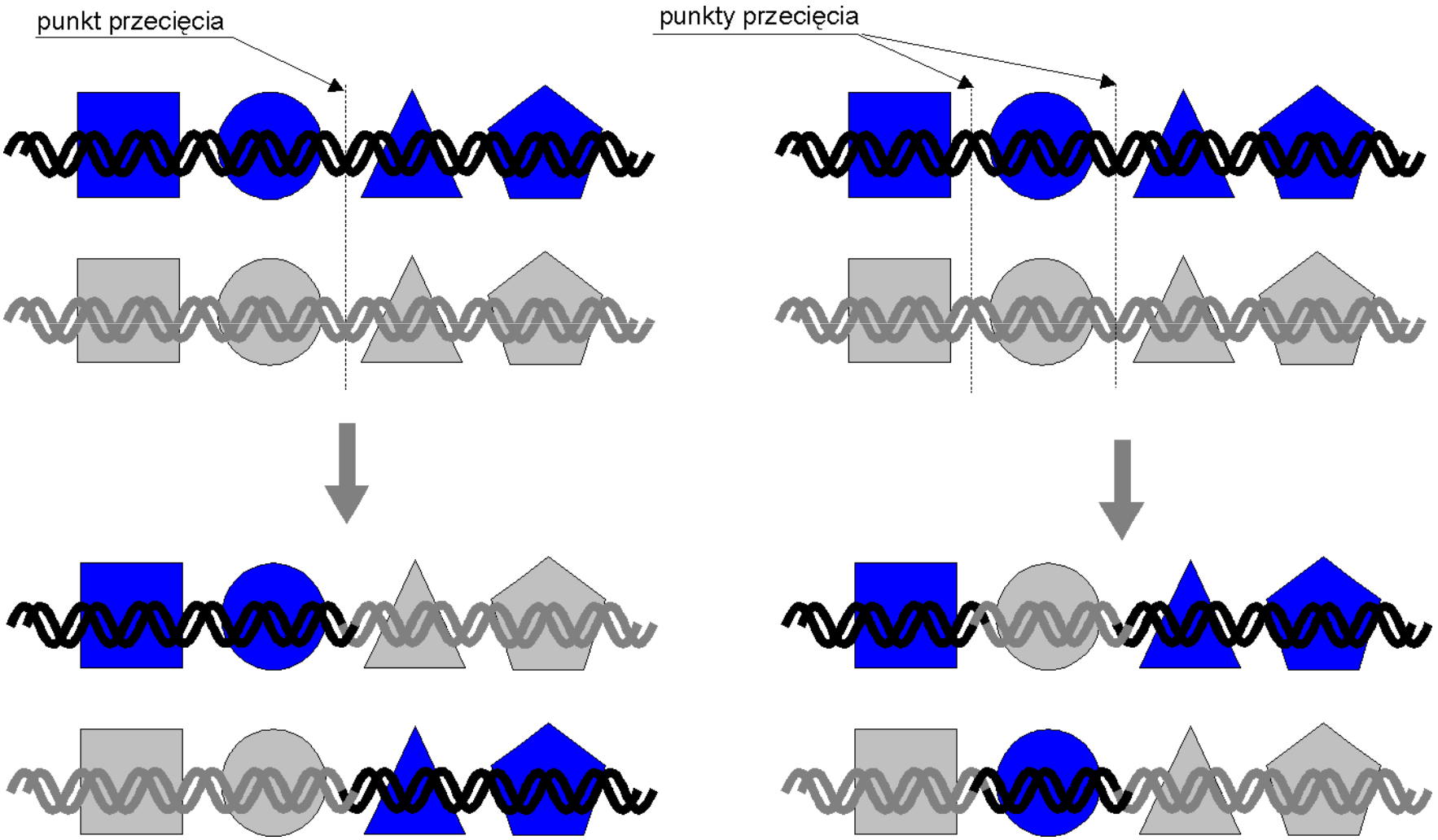
	trzeci element kodonu			
	A	C	G	U
AA	Lizyna (Lys)	Asparagina (Asn)	Lizyna (Lys)	Asparagina (Asn)
AC	Treonina (Thr)	Treonina (Thr)	Treonina (Thr)	Treonina (Thr)
AG	Arginina (Arg)	Seryna (Ser)	Arginina (Arg)	Seryna (Ser)
AU	Izoleucyna (Ile)	Izoleucyna (Ile)	Metionina (Met) - kodon startowy	Izoleucyna (Ile)
CA	Glutamina (Gln)	Histydyna (His)	Glutamina (Gln)	Histydyna (His)
CC	Prolina (Pro)	Prolina (Pro)	Prolina (Pro)	Prolina (Pro)
CG	Arginina (Arg)	Arginina (Arg)	Arginina (Arg)	Arginina (Arg)
CU	Leucyna (Leu)	Leucyna (Leu)	Leucyna (Leu)	Leucyna (Leu)
GA	Kwas glutaminowy (Glu)	Kwas asparaginowy (Asp)	Kwas glutaminowy (Glu)	Kwas asparaginowy (Asp)
GC	Alanina (Ala)	Alanina (Ala)	Alanina (Ala)	Alanina (Ala)
GG	Glicyna (Gly)	Glicyna (Gly)	Glicyna (Gly)	Glicyna (Gly)
GU	Walina (Val)	Walina (Val)	Walina (Val)	Walina (Val)
UA	kodon stopowy	Fenyloalanina (Phe)	kodon stopowy	Fenyloalanina (Phe)
UC	Seryna (Ser)	Seryna (Ser)	Seryna (Ser)	Seryna (Ser)
UG	kodon stopowy	Cysteina (Cys)	Tryptofan (Trp)	Cysteina (Cys)
UU	Leucyna (Leu)	Fenyloalanina (Phe)	Leucyna (Leu)	Fenyloalanina (Phe)

Synteza białaka - translacaja mRNA

Biologiczne tło AG

- **Rekombinacja (krzyżowanie):**
  - poprzedza proces mejozy (tworzenie komórek rozrodczych)
  - wymiana informacji genetycznej zawartej w chromosomach poprzez wymianę odcinków DNA między chromosomami z par homologicznych
  - tworzenie nowych kombinacji genetycznych, czyli sekwencji nukleotydowych

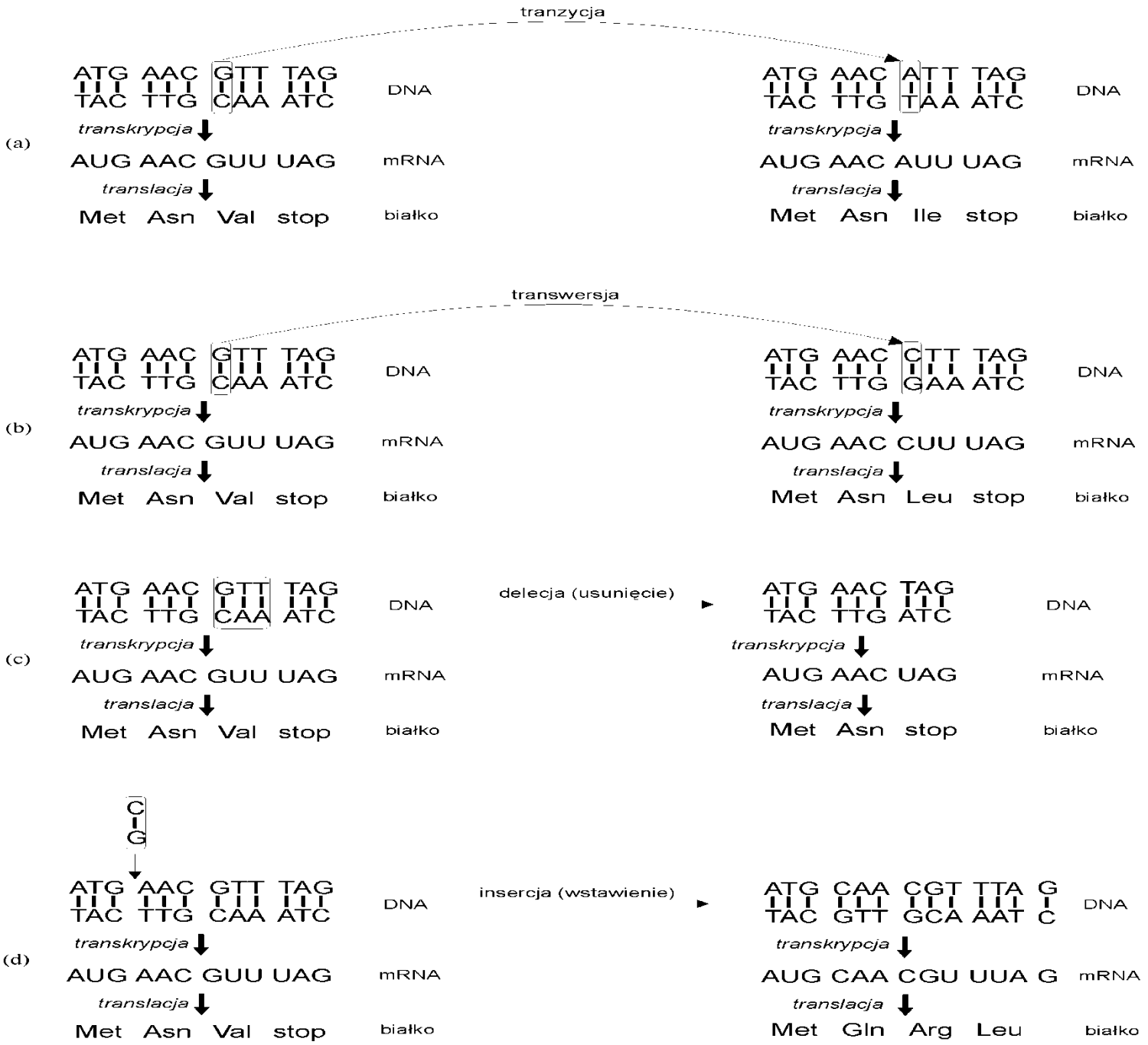
Rekombinacja dla pary łańcuchów DNA z chromosomów homologicznych



Biologiczne tło AG

- W genetyce rozróżnia się cztery typy mutacji:
  - **tranzycja**
  - **transwersja**
  - **delecja**
  - **insercja**

Rodzaje mutacji

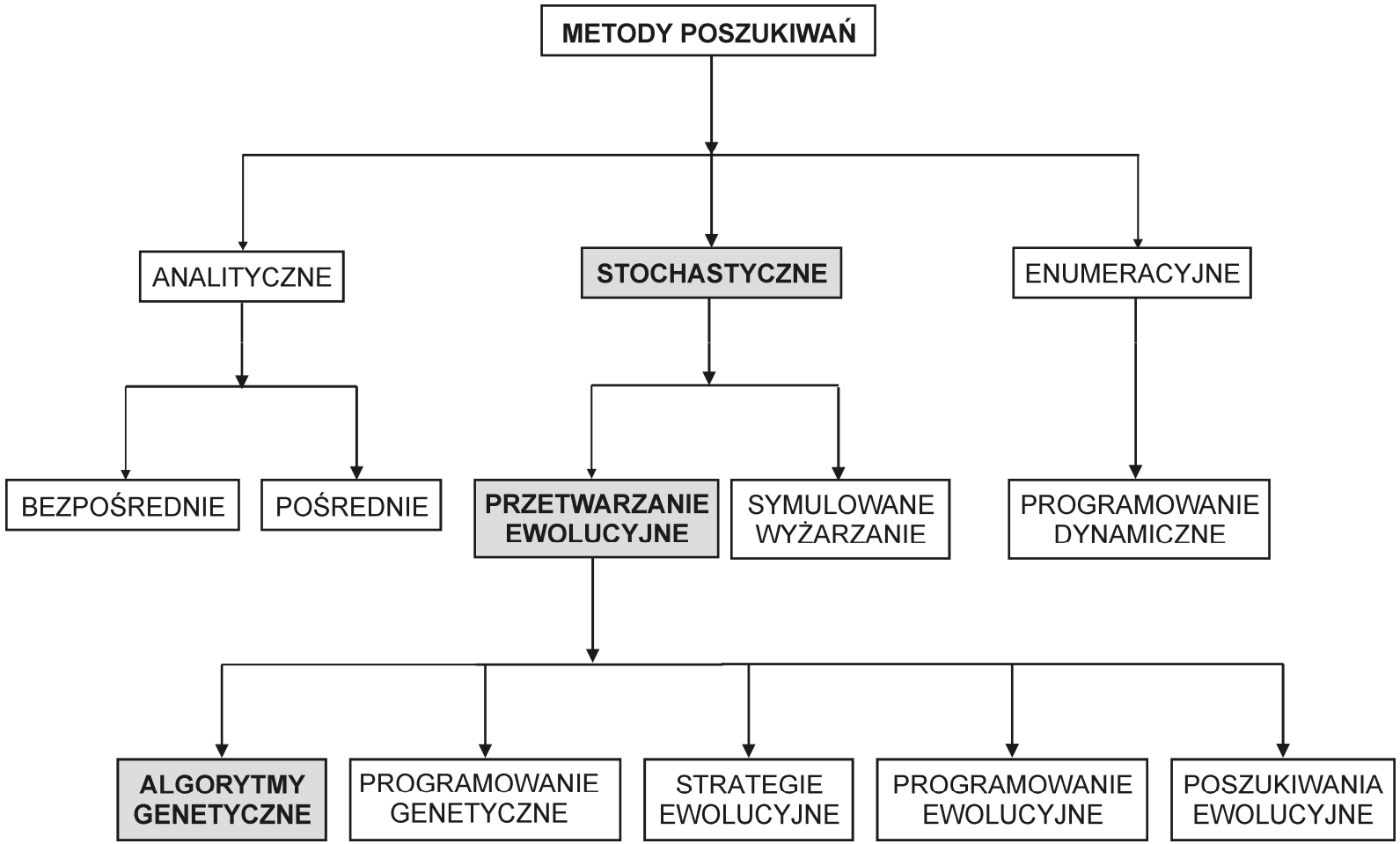


Wprowadzenie

Metody optymalizacji (poszukiwań) można podzielić na trzy klasy:

- **analityczne**
- **stochastyczne**
- **enumeracyjne**

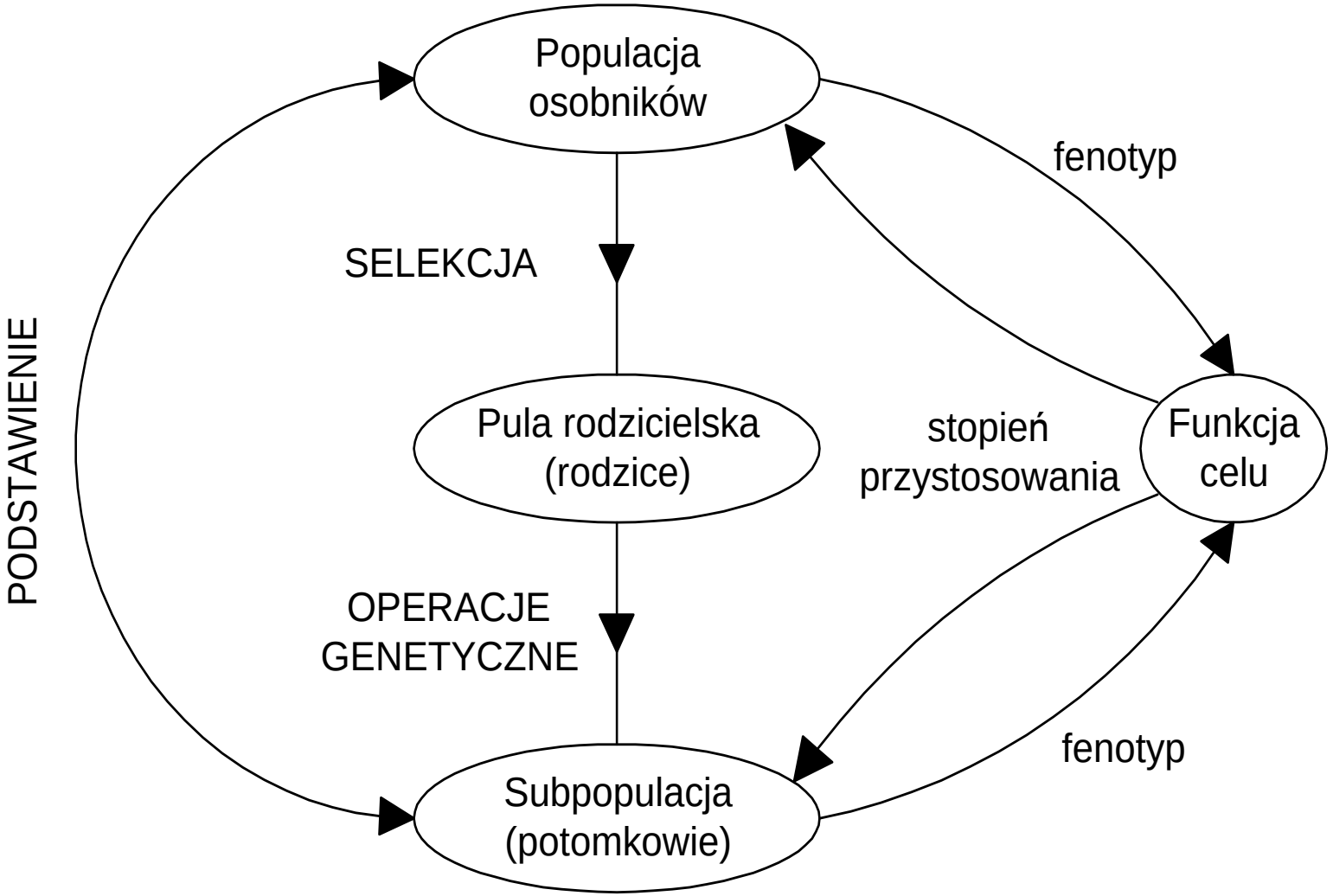
Metody poszukiwań



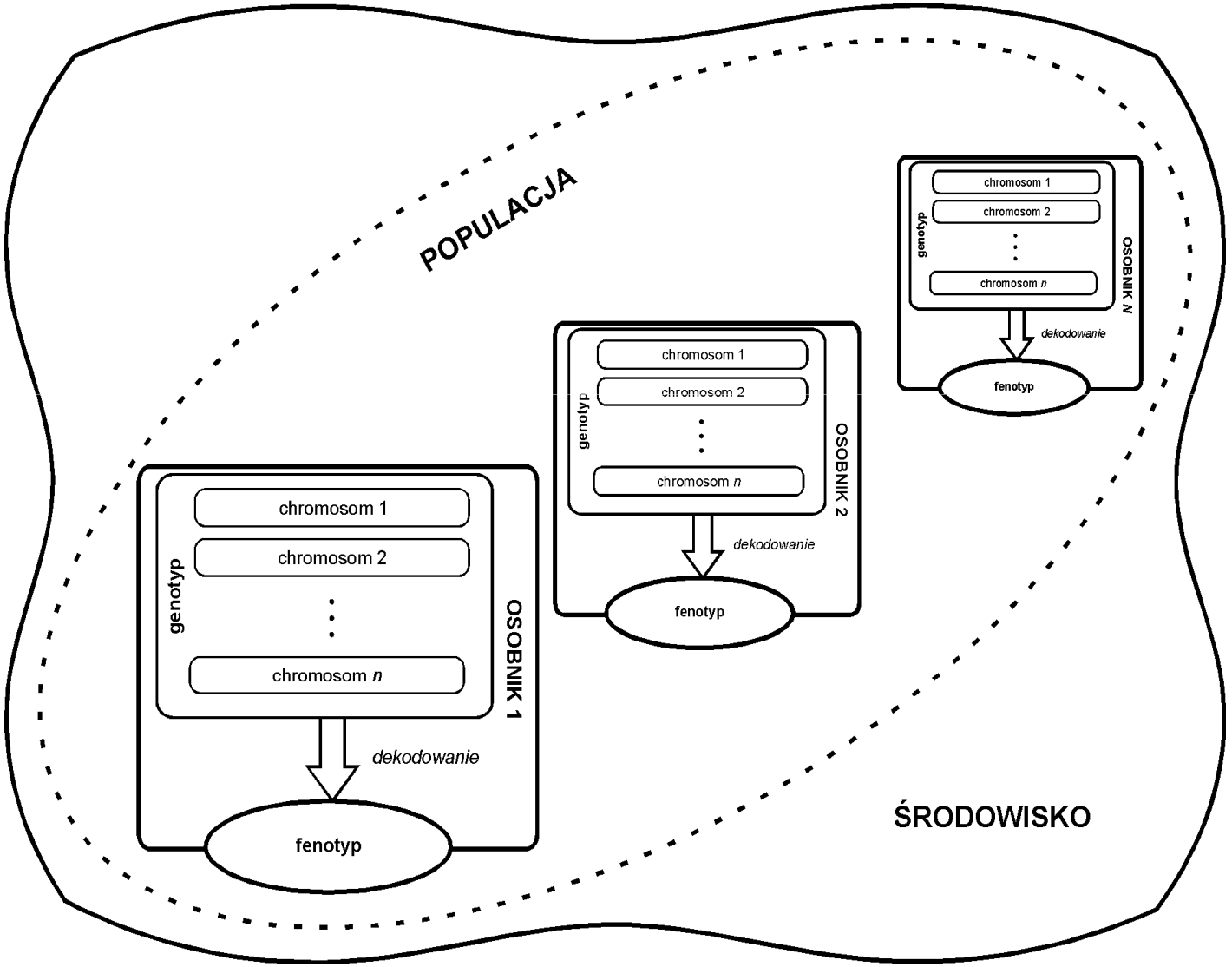
Wprowadzenie

- Podstawowe cechy algorytmów genetycznych:
  - prowadzenie poszukiwań w wielu punktach
  - parametry zadania optymalizacji posiadają postać zakodowaną
  - do wyboru i tworzenia nowych rozwiązań stosuje się reguły ewolucyjno-probabilistyczne
  - wykorzystanie do poszukiwań łatwo obliczalnej wartości funkcji celu
  - wolne są od ograniczeń nakładanych na przestrzeń poszukiwań (np. ciągłość, istnienie pochodnych oraz unimodalność funkcji celu)
  - uniwersalność i prostota implementacyjna

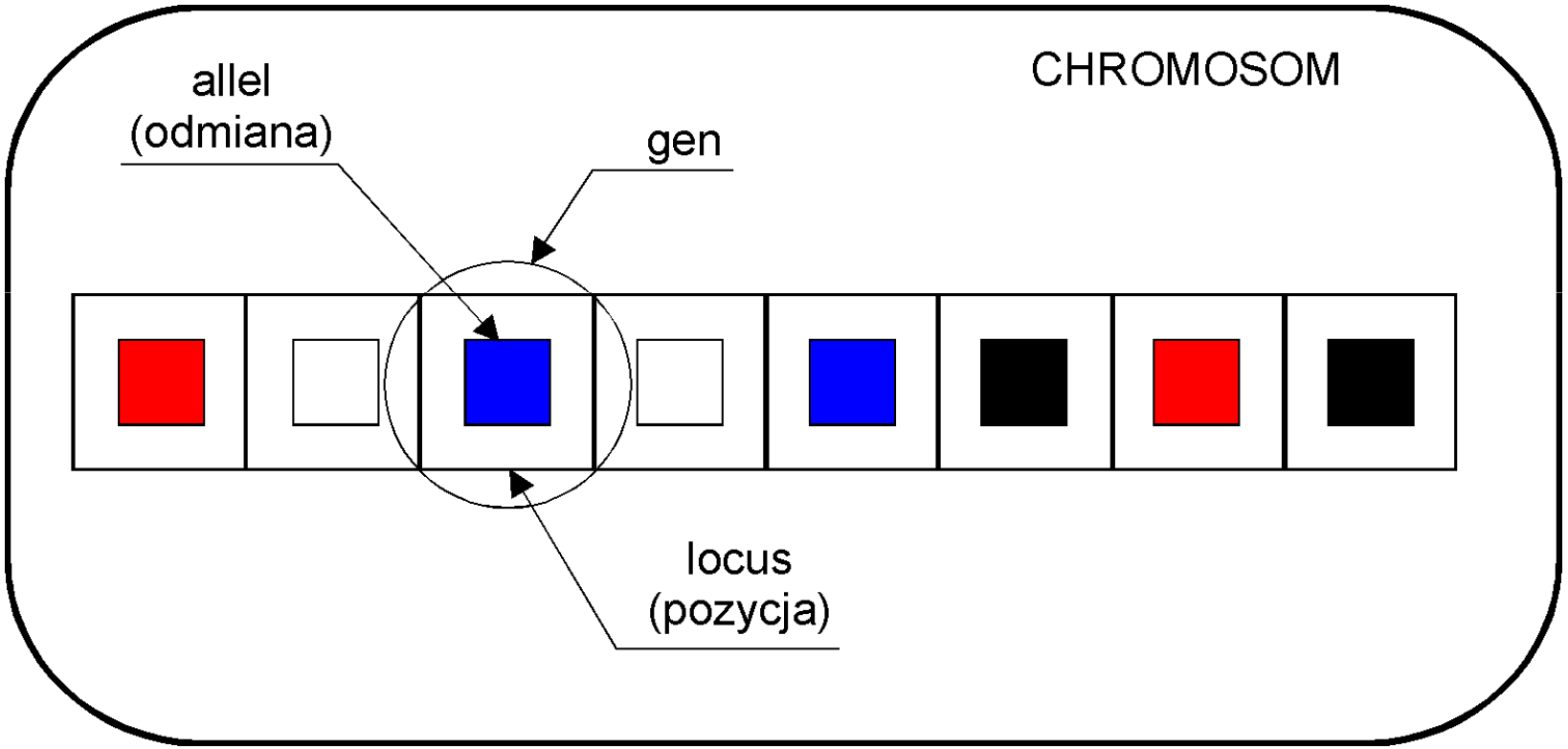
Cykl ewolucyjny AG



Model populacji osobników



Model chromosomu w AG



Podstawowe oznaczenia

- Funkcja przystosowania

$$\forall_{\boldsymbol{x}\in\Re^n} \quad f(\boldsymbol{x}) \geq 0$$

- Genotyp osobnika haploidalnego (chromosom)

$$\boldsymbol{v}_i = \begin{bmatrix} v_{1_i} & v_{2_i} & \dots & v_{n_i} \end{bmatrix}^{\mathrm{T}}, \quad i = 1,2,\dots,N, \quad n \in \aleph$$

- Populacja osobników haploidalnych

$$\boldsymbol{V} = [\boldsymbol{v}_1 \quad \boldsymbol{v}_2 \quad \dots \quad \boldsymbol{v}_N]$$

Podstawowe oznaczenia

- Fenotyp osobnika haploidalnego

$$\boldsymbol{x}_i = \begin{bmatrix} x_{1_i} & x_{2_i} & \dots & x_{n_i} \end{bmatrix}^{\mathrm{T}} \in \Re^n$$

- Stopień przystosowania osobnika haploidalnego

$$\boldsymbol{f}(\boldsymbol{x}_i) \in \Re^m$$

Podstawowe oznaczenia

- Genotyp osobnika diploidalnego

$$\boldsymbol{v}_{di} = \left[\boldsymbol{v}_{1_i} \quad \boldsymbol{v}_{2_i}\right]^{\mathrm{T}}, \quad i = 1, 2, \ldots, N$$

gdzie

$$\boldsymbol{v}_{1_i} = \left[v_{11_i} \quad v_{21_i} \quad \cdots \quad v_{n1_i}\right]^{\mathrm{T}}$$

$$\boldsymbol{v}_{2_i} = \left[v_{12_i} \quad v_{22_i} \quad \cdots \quad v_{n2_i}\right]^{\mathrm{T}}$$

Podstawowe oznaczenia

- Populacja osobników diploidalnych

$$\boldsymbol{V}_d = [\boldsymbol{v}_{d1} \quad \boldsymbol{v}_{d2} \quad \dots \quad \boldsymbol{v}_{dN}]$$

- Fenotyp osobnika diploidalnego

$$\boldsymbol{x}_i = \left[x_{1_i} \quad x_{2_i} \quad \dots \quad x_{n_i}\right]^{\mathrm{T}} \in \Re^n$$

- Stopień przystosowania osobnika diploidalnego

$$\boldsymbol{f}(\boldsymbol{x}_i) \in \Re^m$$

Kodowanie informacji

- Rodzaje kodowania (reprezentacji parametrów):
  - **bialleliczne**
  - **Gray’a**
  - **logarytmiczne**
  - **trialleliczne**
  - **multialleliczne**
  - **wielopoziomowe**

Kodowanie informacji

- **Bialleliczne** - odcinki (współrzędne) chromosomu osobnika są skończonymi sekwencjami genów o allelach ze zbioru {0, 1}

$$\boldsymbol{v}_i = [0100111101 \quad 0100111 \quad 00011011 \quad 10111]^{\mathrm{T}}$$

- **Gray'a** – kodowanie bialleliczne (sąsiednie punkty w przestrzeni poszukiwań różnią się tylko jednym bitem w przestrzeni zakodowanych parametrów

Kodowanie informacji

- **Logarytmiczne** – kodowanie biallelicznе
  - zmniejszenie długości chromosomów dla ogromnych oraz wielowymiarowych przestrzeni poszukiwań
  - bit najstarszy jest bitem znaku funkcji wykładniczej, zaś bit po nim następujący jest bitem znaku wykładnika funkcji wykładniczej
  - pozostałe geny reprezentują wykładnik funkcji wykładniczej

Kodowanie informacji

Przykład 1.

Chromosom z kodowaniem logarytmicznym

$$\boldsymbol{v}_i = [0100111101\ 00011011]^{\mathrm{T}}$$

$$\boldsymbol{v}_i = \left[(-1)^0 e^{(-1)^1 (00111101)} \quad (-1)^0 e^{(-1)^0 (011011)} \right]^{\mathrm{T}}$$

Kodowanie informacji

Przykład 2.

Dla długości chromosomu $m = 10$ i kodowania logarytmicznego można zakodować liczby rzeczywiste z przedziału $\left[e^{-255}, e^{255}\right]$

W przypadku kodowania biallelicznego zakres wynosi $[0, 1023]$

Kodowanie informacji

- **Trialleliczne** - sekwencja genów, których allele należą do {-1, 0, 1}

$$\boldsymbol{v}_{di} = \begin{bmatrix} \boldsymbol{v}_{d1} \\ \boldsymbol{v}_{d2} \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} -1 & -1 & 0 & 1 & 0 & -1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & -1 & 0 & -1 & 1 \end{bmatrix}^{\mathrm{T}}$$

- **Multialleliczne** (**zmiennopozycyjne**) - naturalny sposób kodowania parametrów należących do dziedziny określonej w zbiorze liczb rzeczywistych

$$\boldsymbol{v}_i = \boldsymbol{x}_i \in \Re^n$$

Kodowanie informacji

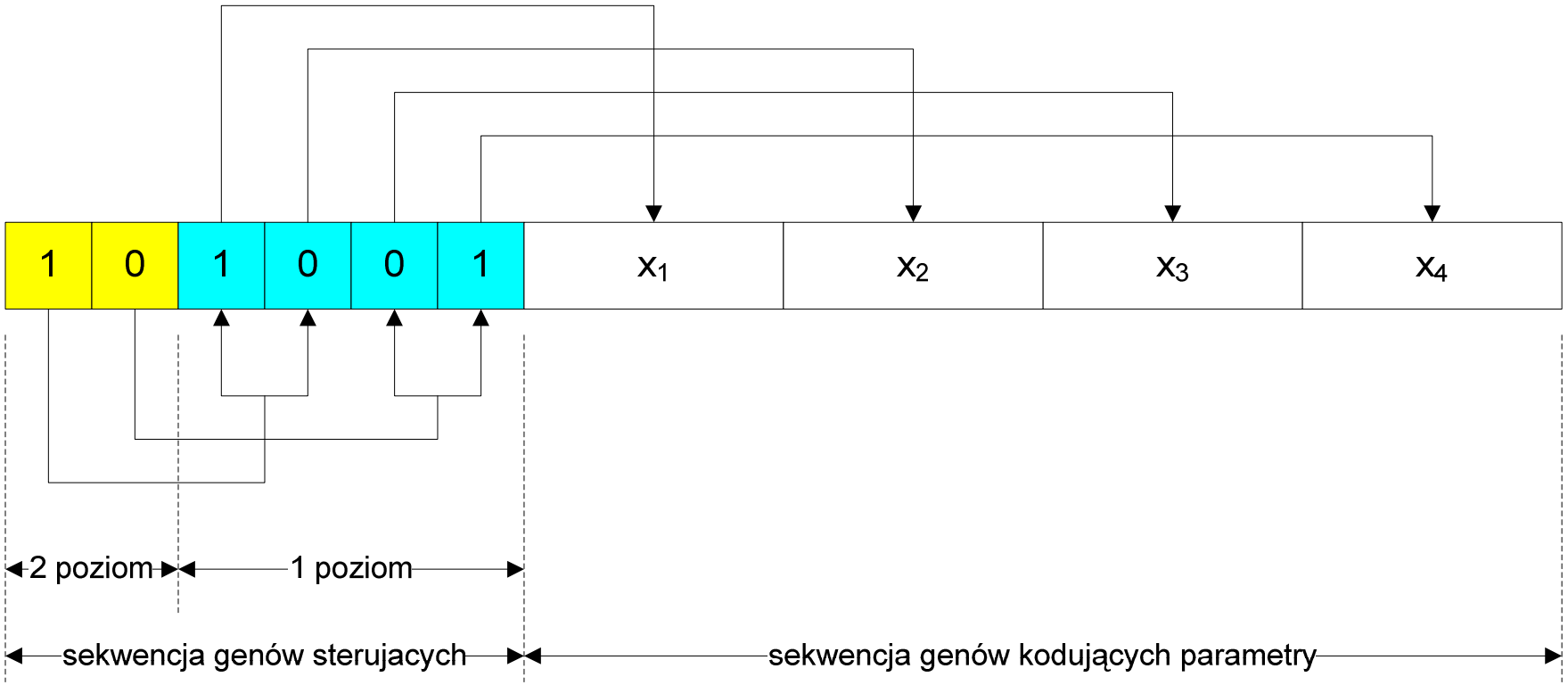
- **Całkowitoliczbowe** stosowane w zadaniach optymalizacji kombinatorycznej

$$\boldsymbol{v}_i = [2 \quad 5 \quad 4 \quad 7 \quad 8 \quad 6 \quad 1 \quad 3]^{\mathrm{T}}$$

- **Wielopoziomowe** - dla problemów parametrycznej i strukturalnej optymalizacji. Genotyp osobnika składa się z dwóch części:
  - sekwencji genów sterujących (strukturalnych)
  - sekwencji genów kodujących parametry zadania optymalizacji (parametryczne)

Kodowanie informacji

- Przykładowe kodowanie dwupoziomowe



Dekodowanie informacji

- **Dla kodowania binarnego**

$$x_{k_i} = \underline{x}_k + \frac{\overline{x}_k - \underline{x}_k}{2^{m_k} - 1} \sum_{j=0}^{m_k} v_{k_{ij}} 2^j$$

gdzie

x_{k_i} - *k*-ta współrzędna wektora parametrów *i*-tego osobnika

m_k - długość sekwencji v_{k_i} kodującej x_{k_i}

$[\underline{x}_k, \overline{x}_k] \subset \Re$ - zakres poszukiwań

Dekodowanie informacji

- **Dla kodowania triallelicznego** (dwie fazy):
  1. Dominacja genetyczna

	-1	0	1
-1	-1	0	1
0	0	0	1
1	1	1	1

  2. Liniowe odwzorowanie

$$x_{k_i} = \underline{x}_k + \frac{\bar{x}_k - \underline{x}_k}{2^{m_k} - 1} \sum_{j=0}^{m_k} \left| v_{k_{ij}} \right| 2^j$$

Ocena przystosowania

- Funkcja celu jest głównym źródłem oceny stopnia przystosowania każdego osobnika
- W algorytmach ewolucyjnych wymaga się, aby funkcja celu była funkcją przystosowania (często niespełniony warunek)

Ocena przystosowania

- Przekształcenie funkcji kosztu $u(\boldsymbol{x}) \in R$ do funkcji przystosowania:

$$f(\boldsymbol{x}) = C_{\max} - u(\boldsymbol{x})$$

gdzie:

$C_{\max}$ - współczynnik nie mniejszy niż maksymalna wartość funkcji uzyskana w aktualnej populacji

Ocena przystosowania

- Przekształcenie funkcji zysku $g(\boldsymbol{x}) \in R$ do funkcji przystosowania:

$$f(\boldsymbol{x}) = g(\boldsymbol{x}) - C_{\min}$$

gdzie:

$C_{\min}$ - współczynnik nie większy niż minimalna wartość funkcji uzyskana w aktualnej populacji

Selekcja osobników

- Naśladuje mechanizm przeżycia w naturze
- Osobnik o najwyższym stopniu przystosowaniu uzyska liczne potomstwo, tj. pomnoży swój materiał genetyczny
- Osobniki o najwyższym stopniu przystosowaniu podnoszą w ten sposób prawdopodobieństwo swojego przeniknięcia („przeżycia”) do następnego pokolenia
- Osobniki o najniższym stopniu przystosowania powinny być eliminowane z opisywanego procesu prokreacji

Selekcja osobników

- Najczęściej stosowane metody selekcji:
  - **metoda proporcjonalna**
  - **metoda ze stochastycznym doborem resztowym**
  - **metoda turniejowa**
  - **metoda rangowa**
  - **metoda progowa**

Nacisk selektywny

- Naturalna tendencja do poprawiania średniego przystosowania populacji podczas cyklu ewolucyjnego

- Algorytm genetyczny posiada tym większy nacisk selektywny, im większa jest wartość oczekiwana liczby kopii lepszego osobnika w porównaniu do osobnika gorszego

Metoda proporcjonalna

- Konstruowanie tzw. ruletki, której sektory kątowe są proporcjonalne do wartości funkcji przystosowania osobników
- *N*–krotne (*N* – liczba osobników w populacji) ustalenie pozycji zakręcanego koła ruletki
- Lepiej przystosowane osobniki to proporcjonalnie większe sektory kątowe na tarczy ruletki (większa szansa na wprowadzenie większej liczby potomków do następnego pokolenia populacji

Metoda proporcjonalna

1. Wyznacz prawdopodobieństwa wyboru osobników z populacji w postaci względnego przystosowania

$$p_S(\boldsymbol{x}_i) = f(\boldsymbol{x}_i) / \sum_{i=1}^{N} f(\boldsymbol{x}_i), \quad i = 1, 2, \ldots, N$$

2. Wyznacz dystrybuantę dla ustalonej sekwencji osobników

$$q(\boldsymbol{x}_i) = \sum_{j=1}^{i} p_S(\boldsymbol{x}_j)$$

Metoda proporcjonalna

3. *N*-krotnie ‘zakręć kołem ruletki’ poprzez:

   – wygenerowanie losowej liczby $r \in [0,1]$

   – ustalenie pierwszego napotkanego osobnika w populacji, dla którego spełniony jest warunek

$$r \leq q(\boldsymbol{x}_i)$$

Metoda proporcjonalna

Przykład 3. Funkcja kryterialna

$$f(x_1, x_2, x_3) = 1300 - x_1^2 - x_2^2 - x_3^2$$

określona w dziedzinie

$$x_1 \in [0,15] \qquad x_2 \in [0,7] \qquad x_3 \in [0,31]$$

Populacja składa się z 4 osobników haploidalnych

$$\boldsymbol{v}_i = \begin{bmatrix} v_{1_i} & v_{2_i} & v_{3_i} \end{bmatrix}^{\mathrm{T}}, \quad i = 1,2,3,4$$

odcinki chromosomu kodowane binarnie odpowiednio na 4, 3 i 5 bitach

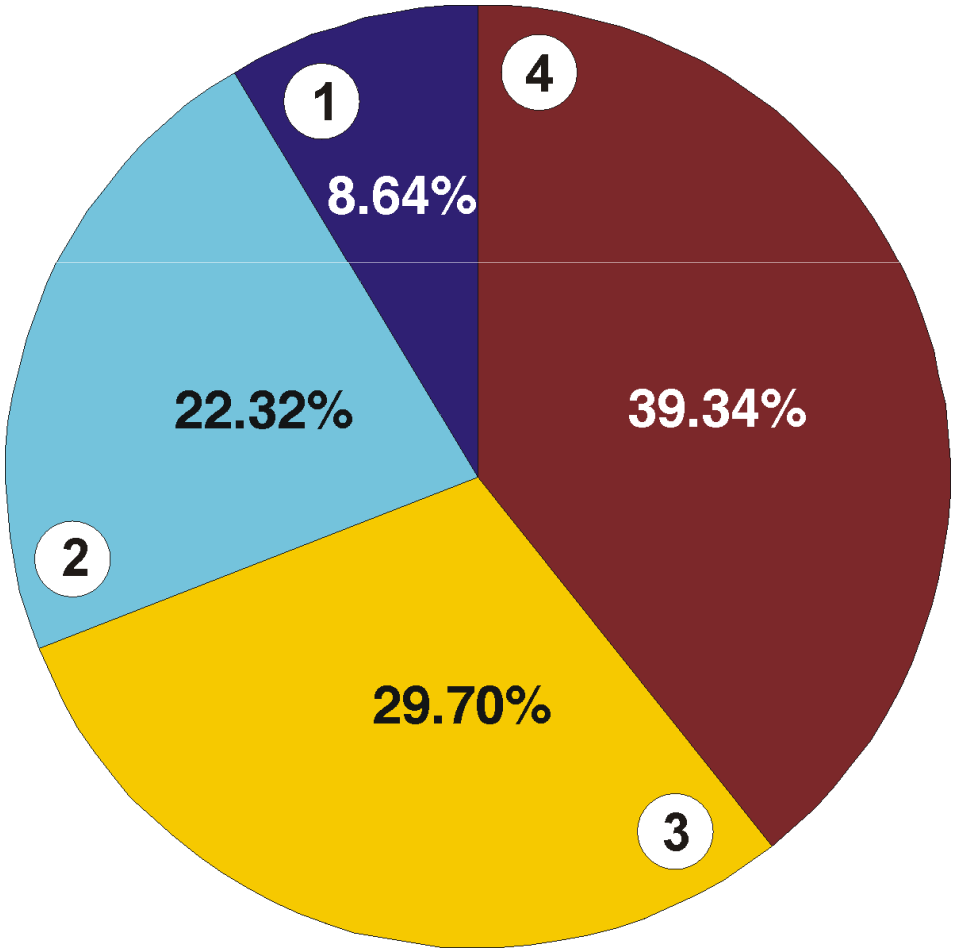
Metoda proporcjonalna

Genotyp, fenotyp i przystosowanie osobnika

Lp.	genotyp	fenotyp	przystosowanie		dystrybuanta
			bezwzględne	względne	
1	$[0110\ \ 110\ \ 11110]^T$	$[6\ \ 6\ \ 31]^T$	267	0.0864	0.0864
2	$[1100\ \ 101\ \ 10101]^T$	$[12\ \ 5\ \ 21]^T$	690	0.2232	0.3096
3	$[0111\ \ 011\ \ 10010]^T$	$[7\ \ 3\ \ 18]^T$	918	0.2970	0.6066
4	$[0100\ \ 010\ \ 01000]^T$	$[4\ \ 2\ \ 8]^T$	1216	0.3934	1.0000
$\sum_{i=1}^{4} f(x_i)$ = 3091					

Metoda proporcjonalna

Koło ruletki dla rozważanego przykładu



Metoda stochastycznego doboru resztowego

1. Wyznacz oczekiwaną liczbę kopii osobników w populacji

$$e(\boldsymbol{x}_i) = N \cdot f(\boldsymbol{x}_i) / \sum_{i=1}^{N} f(\boldsymbol{x}_i), \quad i = 1, 2, \ldots, N$$

2. Skopiuj osobników do puli rodzicielskiej na podstawie całkowitych częściach liczby $e(\boldsymbol{x}_i)$, oraz określ liczbę ‘wakatów’ $\tilde{N} = N - N_{\text{int}}$

Metoda stochastycznego doboru resztowego

3. Wyznacz ‘dystrybuantę’ osobników wg resztowej części $e(\boldsymbol{x}_i)$

$$q(\boldsymbol{x}_i) = \sum_{j=1}^{i} \left\{ e(\boldsymbol{x}_j) - \left\lfloor e(\boldsymbol{x}_j) \right\rfloor \right\}$$

4. $\tilde{N}$ -krotnie ‘zakręć kołem ruletki’ poprzez:
   - wygenerowanie losowej liczby $r \in [0,1]$
   - ustalenie pierwszego napotkanego osobnika spełniającego warunek

$$r \leq \frac{q(\boldsymbol{x}_i)}{q(\boldsymbol{x}_N)}$$

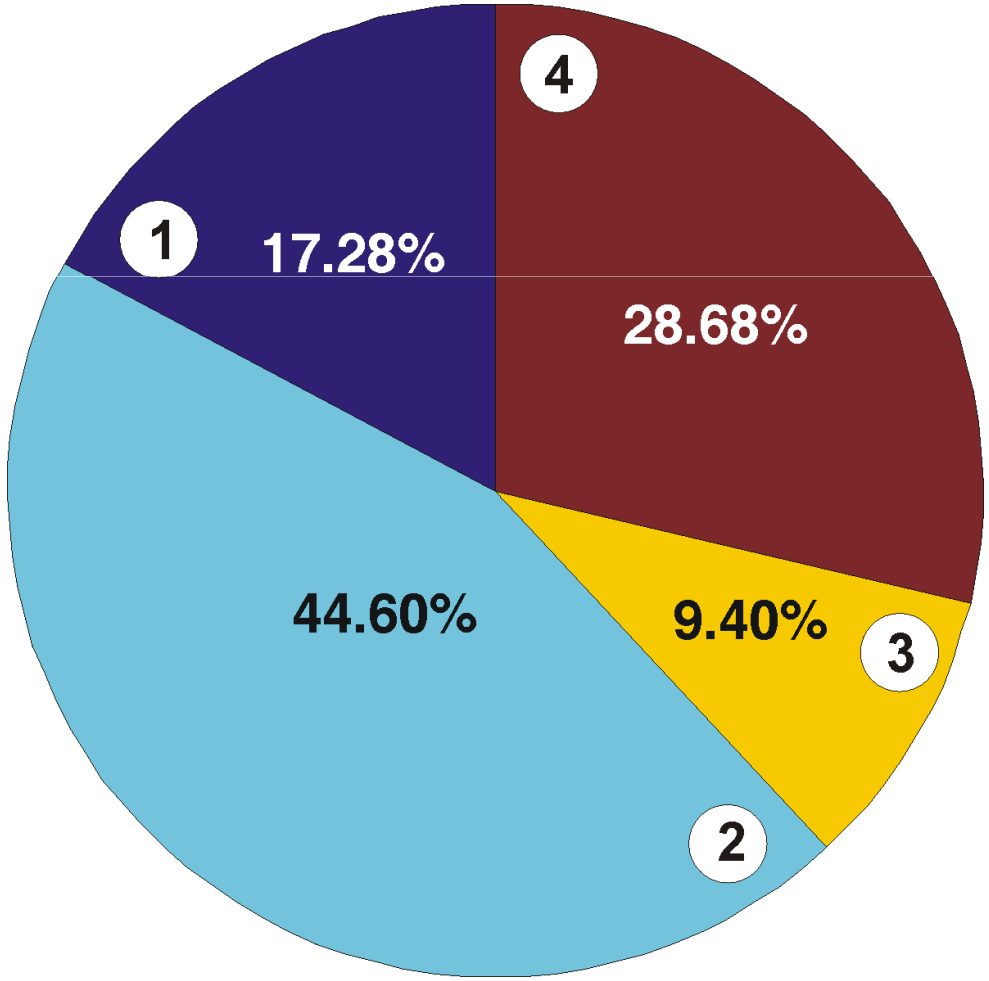
Metoda stochastycznego doboru resztowego

Dla przykładu 3

v_i	$f(x_i)$	$\lfloor e(x_i) \rfloor$	$e(x_i) - \lfloor e(x_i) \rfloor$	$q(x_i)$	$\frac{q(x_i)}{q(x_4)}$	$p_s = \frac{e(x_i) - \lfloor e(x_i) \rfloor}{q(x_4)}$
v_1	267	0	0.3456	0.3456	0.1728	0.1728
v_2	690	0	0.8928	1.2384	0.6192	0.4460
v_3	918	1	0.1880	1.4264	0.7132	0.0940
v_4	1216	1	0.5736	2.0000	1.0000	0.2868
$\sum_{i=1}^{4} f(x_i) = 3091$						

Metoda stochastycznego doboru resztowego

Koło ruletki dla rozważanego przykładu



Metoda turniejowa

- Sposób selekcji został skojarzony z procesem rywalizacji dwóch osobników, z których wygrywa lepiej przystosowany
- Liczba osobników uczestniczących w turnieju decyduje o tzw. nacisku selekcyjnym (ang. *selection pressure*)
- Jeżeli liczba uczestników turnieju jest duża, wówczas słabsze osobniki mają mniejszą szansę na wybranie do puli rodzicielskiej

Metoda turniejowa

1. Wybierz losowo z rozkładem równomiernym *r* – osobników z populacji *N* osobników (*r*<*N*)
2. Przeprowadź turniej dla *r*-osobników poprzez:
   - wyznaczenie zwycięzcy spełniającego warunek

$$\exists_{i} \forall_{j \neq i} f(\boldsymbol{x}_i) \geq f(\boldsymbol{x}_j)$$

   - skopiowanie zwycięzcy turnieju do puli rodzicielskiej
3. powróć do kroku 1 i powtórz cały cykl $N-1$ krotnie

Metoda turniejowa

- Proces wyboru osobników może być przeprowadzany w dwojaki sposób poprzez:
  - losowanie ze zwracaniem
  - losowanie bez zwracania
- Gdy warunek zwycięzcy nierozstrzygnięty wówczas zwycięzca jest wybierany losowo
- Jeżeli grupa turniejowa składa się z dwóch osobników to selekcja nazywana jest **turniejem binarnym**

Metoda rangowa

- Stopień przystosowania jest reprezentowany przez rangę, która nadawana jest w procesie sortowania osobników w populacji
- Sortowanie polega na uszeregowaniu osobników w populacji według nierosnących wartości funkcji przystosowania
- Dwa sposoby przypisywania rang:
  - ranga jest równa indeksowi (pozycji) w uszeregowanej populacji
  - ranga identyczna dla osobników identycznych

Metoda rangowa

1. Posortuj osobniki w populacji według nierosnących wartości funkcji przystosowania
2. Nadaj każdemu osobnikowi rangę $r(\boldsymbol{x}_i)$
   - przypisanie kolejnego indeksu (pozycji) w uszeregowanej populacji albo
   - przypisanie jednakowej pozycji dla osobników o identycznym stopniu przystosowania

Metoda rangowa

3. Wyznacz prawdopodobieństwa wyboru osobników wg odwzorowania:
   – liniowego

$$p_s(\boldsymbol{x}_i) = a + k\left(1 - \frac{r(\boldsymbol{x}_i)}{r_{\max}}\right)$$

   – lub nieliniowego

$$p_s(\boldsymbol{x}_i) = a + k\left(r_{\max} - r(\boldsymbol{x}_i)\right)^b$$

gdzie

$$r_{\max} = \max_i \{r(\boldsymbol{x}_i)\}$$

Metoda rangowa

a, k, b – parametry dobieranie arbitralnie aby spełnione były następujące warunki:

- $\sum_{i=1}^{N} p_s(\boldsymbol{x}_i) = 1\ ,\ \ 0 \le p_s(\boldsymbol{x}_i) \le 1$
- $jeżeli \quad r(\boldsymbol{x}_i) \ge r(\boldsymbol{x}_j) \quad to \quad p_s(\boldsymbol{x}_i) \le p_s(\boldsymbol{x}_j)$

4. Na podstawie prawdopodobieństw wyboru przeprowadź selekcje według metody proporcjonalnej lub metody ze stochastycznym doborem resztowym

Metoda progowa

- Sortowanie osobników i nadawanie rang
- Prawdopodobieństwo wyboru osobnika

$$p_s(\boldsymbol{x}_i) = \begin{cases} \dfrac{1}{\gamma} & dla \quad 0 \le r(\boldsymbol{x}_i) \le \gamma \\ 0 & dla \quad r(\boldsymbol{x}_i) > \gamma \end{cases}$$

gdzie

γ - parametr sterujący naciskiem selektywnym

- Zmniejszanie wartości γ zwiększa nacisk selektywny

Operacje genetyczne

- Operacje genetyczne wykonywane na osobnikach rodzicielskich:
  - **krzyżowanie** oraz
  - **mutacja**.
- Chromosom osobnika traktowany jest jako pojedyncza sekwencja kodowa
- Dla kodowania multiallelicznego operacje genetyczne wykonuje się odrębnie dla poszczególnych współrzędnych (genów)

Krzyżowanie

- **Krzyżowanie** (ang. *crossover*) analogiczna do rekombinacji pomiędzy nićmi DNA chromosomów z par homologicznych
- W GA krzyżowanie polega na wymianie materiału genetycznego (części chromosomu) między dwoma rodzicami
- W wyniku tego uzyskuje się dwóch potomków stanowiących nowe rozwiązania

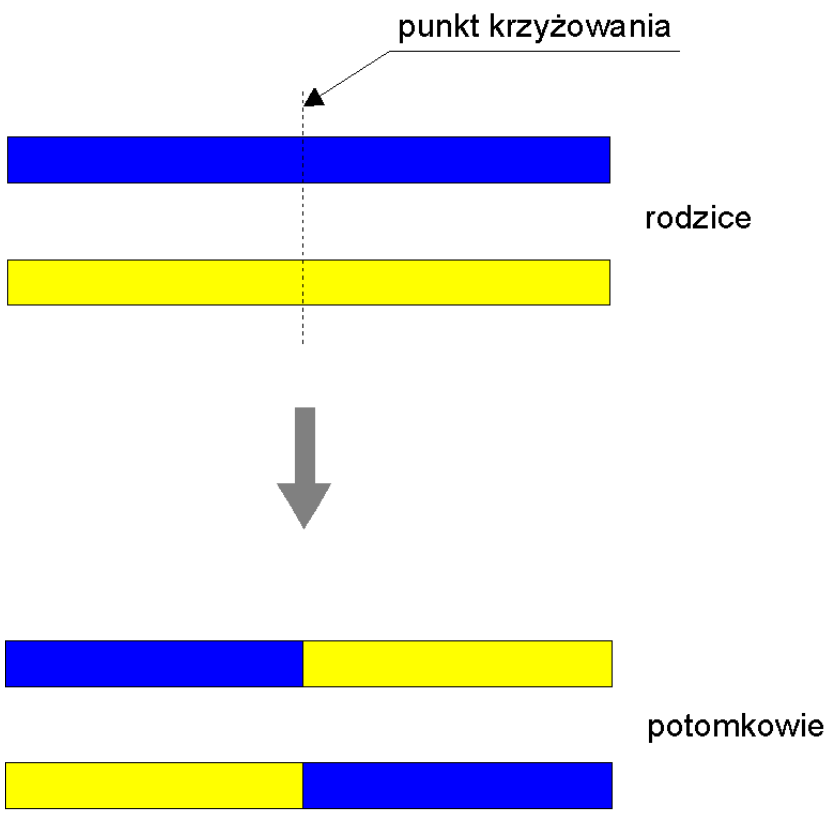
Krzyżowanie

- Krzyżowanie przebiega w dwóch etapach:
  1. Kojarzenie w sposób losowy osobników z puli rodzicielskiej w pary
  2. Każda para rodziców przechodzi proces wymiany materiału genetycznego (tj. wymiany fragmentów sekwencji kodowych pomiędzy chromosomami osobników rodzicielskich)
- Krzyżowanie zachodzi z pewnym prawdopodobieństwem $p_c \in [0.6, 1]$ nazywanym **prawdopodobieństwem krzyżowania**

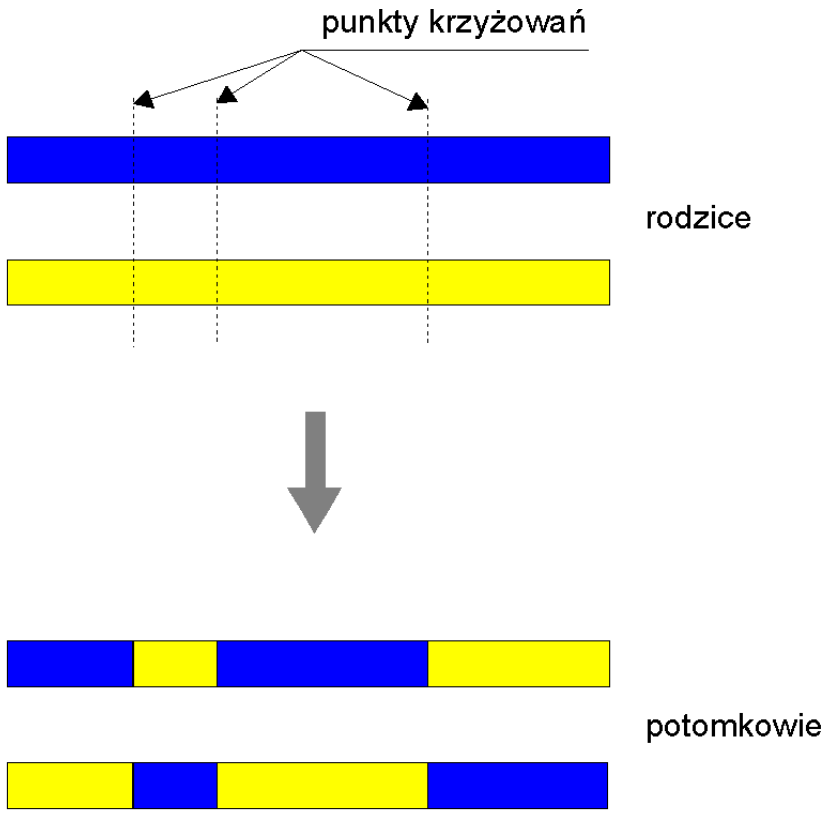
Podstawowe sposoby krzyżowania

- Najczęściej stosowanym sposobem krzyżowania jest:
  - **krzyżowanie jednopunktowe**,
  - **krzyżowanie wielopunktowe** (zwykle dwu- lub trójpunktowe) oraz
  - **krzyżowanie jednorodne**.

Podstawowe sposoby krzyżowania



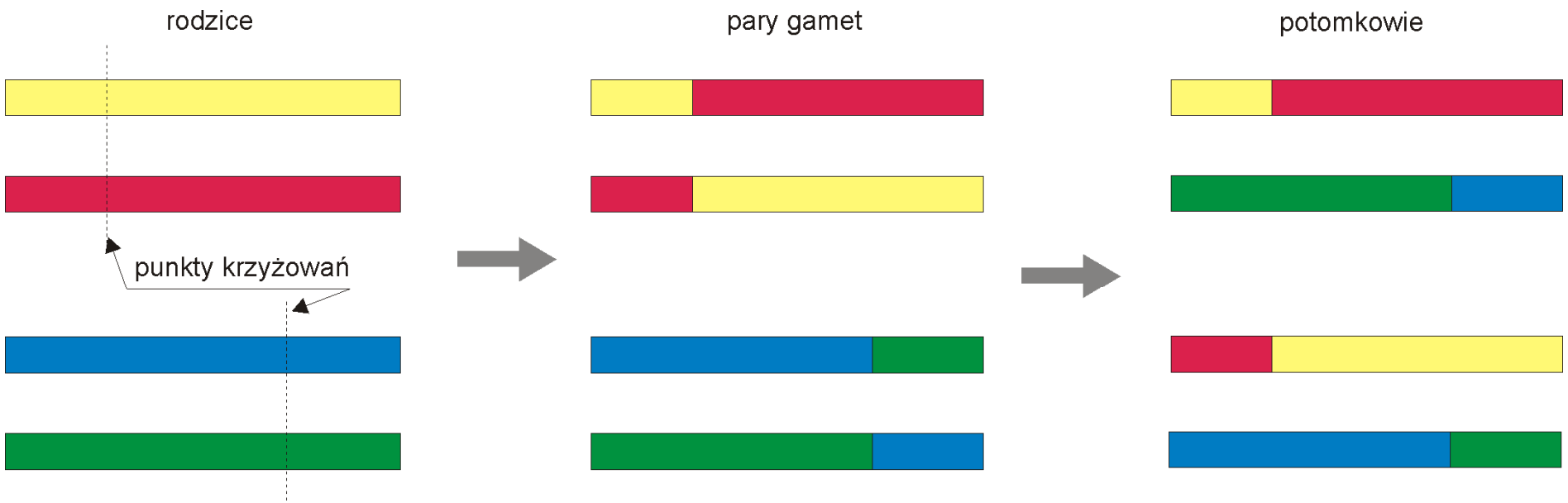
jednopunktowe



trójpunktowe

Podstawowe sposoby krzyżowania

Jednopunktowe dla osobników diploidalnych



Podstawowe sposoby krzyżowania

Przykład 4. Krzyżowanie jednopunktowe dla osobników kodowanych binarnie

$$\boldsymbol{v}_r = [0100 \quad 11110 \quad 0100111]^{\mathrm{T}}$$

$$\boldsymbol{v}_s = [0100 \quad 00011 \quad 0000000]^{\mathrm{T}}$$

Punkt krzyżowania pomiędzy 10 a 11 bitem

$$\boldsymbol{v}_r' = [0100 \quad 11011 \quad 0000000]^{\mathrm{T}}$$

$$\boldsymbol{v}_s' = [0100 \quad 00110 \quad 0100111]^{\mathrm{T}}$$

Podstawowe sposoby krzyżowania

- **Krzyżowanie równomierne** (jednolite) (ang. *uniform crossover*)
- Przeprowadzane wg binarnego wzorca krzyżowania generowanego losowo dla każdej pary rodzicielskich osobników
- Allele równe 1 z takiego wzorca decydują o wymianie pomiędzy rodzicami genów z pozycji im odpowiadającym

Podstawowe sposoby krzyżowania

Przykład 5. Krzyżowanie jednorodne

$$\boldsymbol{v}_r = [0101 \quad 10110 \quad 0100111]^{\mathrm{T}}$$
$$\boldsymbol{v}_s = [1110 \quad 00011 \quad 0001000]^{\mathrm{T}}$$

Wylosowany wzorzec krzyżowania

$$\boldsymbol{w} = [0110 \quad 10100 \quad 0110101]^{\mathrm{T}}$$

Uzyskani potomkowie

$$\boldsymbol{v}_r' = [0100 \quad 11011 \quad 0000000]^{\mathrm{T}}$$
$$\boldsymbol{v}_s' = [0100 \quad 00110 \quad 0100111]^{\mathrm{T}}$$

Podstawowe sposoby krzyżowania

- Dla reprezentacji zmiennopozycyjnej stosuje się:
  - **krzyżowanie arytmetyczne** lub
  - **krzyżowanie mieszane**

- Utworzeni potomkowie stanowią rozwiązania dopuszczalne, tzn. należą do dziedziny określonej na podstawie zakresu parametrów i innych liniowych ograniczeń nałożonych na te parametry

Podstawowe sposoby krzyżowania

- **Krzyżowanie arytmetyczne** - unormowana liniowa kombinacji dwóch wektorów

$$\boldsymbol{v}_r = \begin{bmatrix} v_{1_r} & v_{2_r} & \cdots & v_{n_r} \end{bmatrix}^{\mathrm{T}}$$

$$\boldsymbol{v}_s = \begin{bmatrix} v_{1_s} & v_{2_s} & \cdots & v_{n_s} \end{bmatrix}^{\mathrm{T}}$$

- Potomkowie wyznaczani z poniższych formuł

$$\boldsymbol{v}_r^{'} = a\boldsymbol{v}_r + (1-a)\boldsymbol{v}_s \,, \qquad \boldsymbol{v}_s^{'} = a\boldsymbol{v}_s + (1-a)\boldsymbol{v}_r$$

gdzie

$a \in [0,1]$ - wybierana losowo dla pary rodziców

Podstawowe sposoby krzyżowania

- **Krzyżowanie mieszane (proste)** - połączenie krzyżowania jednopunktowego z krzyżowaniem arytmetycznym

$$\boldsymbol{v}_r = \begin{bmatrix} v_{1_r} & v_{2_r} & \dots & v_{n_r} \end{bmatrix}^{\mathrm{T}}$$

$$\boldsymbol{v}_s = \begin{bmatrix} v_{1_s} & v_{2_s} & \dots & v_{n_s} \end{bmatrix}^{\mathrm{T}}$$

- Potomkowie wyznaczani z poniższych formuł

$$\boldsymbol{v}_r' = \begin{bmatrix} v_{1_r} & \dots & v_{k_r} & av_{(k+1)_s} + (1-a)v_{(k+1)_r} & \dots & av_{n_s} + (1-a)v_{n_r} \end{bmatrix}^{\mathrm{T}}$$

$$\boldsymbol{v}_s' = \begin{bmatrix} v_{1_s} & \dots & v_{k_s} & av_{(k+1)_r} + (1-a)v_{(k+1)_s} & \dots & av_{n_r} + (1-a)v_{n_s} \end{bmatrix}^{\mathrm{T}}$$

Podstawowe sposoby krzyżowania

- **Krzyżowanie wieloosobnicze -** do utworzenia potomków wybieranych jest losowo więcej niż dwóch rodziców
- Dla reprezentacji fenotypowej

$$\mathbf{v}_i' = \sum_{j=1}^{k} a_j \mathbf{v}_j$$

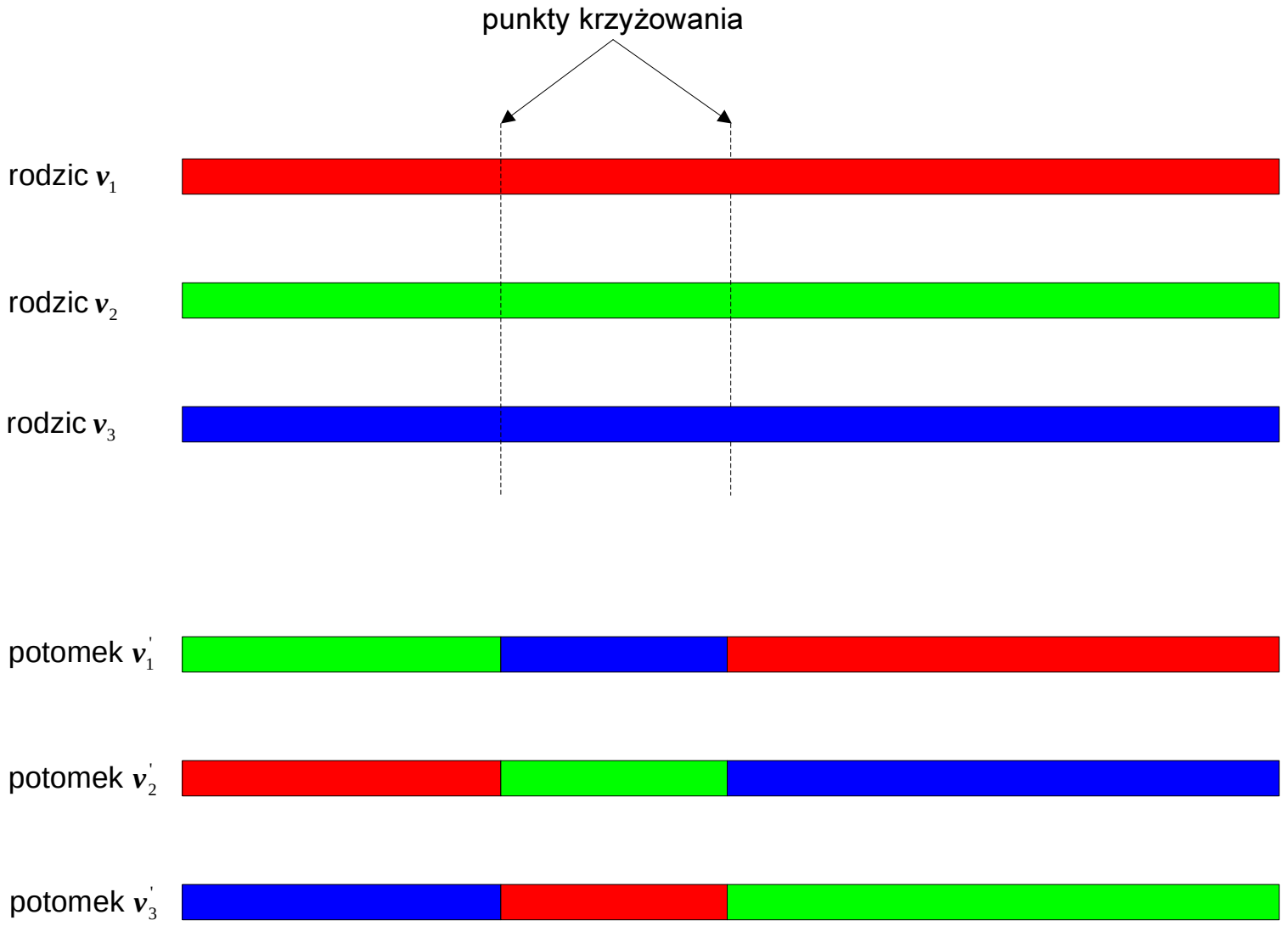
gdzie

a_j - losowe liczby z rozkładem równomiernym,

przy warunku $\sum_{j=1}^{k} a_j = 1$

Podstawowe sposoby krzyżowania

Trójosobnicze dwupunktowe



Podstawowe sposoby krzyżowania

- Dla kodowania całkowitoliczbowego stosuje się zwykle krzyżowanie:
  - **z częściowym odwzorowaniem** (ang. *partialy-mapped crossover*, **PMX**)
  - **z zachowaniem porządku** (ang. *order crossover*, **OX**)
  - **cykliczne** (ang. *cycle crossover*, **CX**)

Podstawowe sposoby krzyżowania

- **Krzyżowanie PMX** polega na wylosowaniu dwóch punktów krzyżowania dla wybranej losowo pary rodziców
- Wybrana sekwencja genów ulega zamianie
- Sekwencje pozostałych genów są tak dobierane, aby zachować porządek genów i ich pozycje w obu powstających potomkach

Podstawowe sposoby krzyżowania

Przykład 6. Krzyżowanie PMX

$$\boldsymbol{v}_r = [4 \quad 2 \quad | \quad 8 \quad 7 \quad 5 \quad 9 \quad | \quad 1 \quad 3 \quad 6]^{\mathrm{T}}$$

$$\boldsymbol{v}_s = [1 \quad 9 \quad | \quad 7 \quad 5 \quad 4 \quad 6 \quad | \quad 8 \quad 2 \quad 3]^{\mathrm{T}}$$

- *Wymiana sekwencji genów pomiędzy punktami krzyżowań*
- *Pozostałe pozycje nieokreślone*

$$\boldsymbol{v}_r' = [? \quad ? \quad | \quad 7 \quad 5 \quad 4 \quad 6 \quad | \quad ? \quad ? \quad ?]^{\mathrm{T}}$$

$$\boldsymbol{v}_s' = [? \quad ? \quad | \quad 8 \quad 7 \quad 5 \quad 9 \quad | \quad ? \quad ? \quad ?]^{\mathrm{T}}$$

Podstawowe sposoby krzyżowania

- *Uzupełnianie pozycji nieokreślonych genami, dla których nie występuje konflikt*
- *Niedopuszczalne powtórzenia alleli w osobniku potomnym (każde rozwiązanie jest permutacją 9-elementowego ciągu cyfr od 1 do 9*

$$\boldsymbol{v}_r' = [?\quad 2\quad |\quad 7\quad 5\quad 4\quad 6\quad |\quad 1\quad 3\quad ?]^{\mathrm{T}}$$

$$\boldsymbol{v}_s' = [1\quad ?\quad |\quad 8\quad 7\quad 5\quad 9\quad |\quad ?\quad 2\quad 3]^{\mathrm{T}}$$

Podstawowe sposoby krzyżowania

- *Wypełnianie pozycji nieokreślonych wg ciągu odwzorowań wymiany sekwencji genów (określony na podstawie punktów krzyżowań)*

$$8 \leftrightarrow 7, 7 \leftrightarrow 5, 5 \leftrightarrow 4, 9 \leftrightarrow 6$$

- *Ostatecznie uzyskani potomkowie*

$$\boldsymbol{v}_r^{'} = [8 \quad 2 \quad | \quad 7 \quad 5 \quad 4 \quad 6 \quad | \quad 1 \quad 3 \quad 9]^{\mathrm{T}}$$
$$\boldsymbol{v}_s^{'} = [1 \quad 6 \quad | \quad 8 \quad 7 \quad 5 \quad 9 \quad | \quad 4 \quad 2 \quad 3]^{\mathrm{T}}$$

Podstawowe sposoby krzyżowania

- **Krzyżowanie OX** polega na skopiowaniu losowych sekwencji genów ustalonych na podstawie dwóch punktów krzyżowań

- Nieokreślone pozycje w tworzonych nowych rozwiązaniach wyznaczane są na podstawie wzajemnego uporządkowania genów jednego z osobników rodzicielskich

Podstawowe sposoby krzyżowania

Przykład 7. Krzyżowanie OX

$$\boldsymbol{v}_r = [4 \quad 2 \quad | \quad 8 \quad 7 \quad 5 \quad 9 \quad | \quad 1 \quad 3 \quad 6]^{\mathrm{T}}$$

$$\boldsymbol{v}_s = [1 \quad 9 \quad | \quad 7 \quad 5 \quad 4 \quad 6 \quad | \quad 8 \quad 2 \quad 3]^{\mathrm{T}}$$

- *Kopiowanie sekwencji genów pomiędzy punktami krzyżowań*
- *Pozostałe pozycje nieokreślone*

$$\boldsymbol{v}_r' = [? \quad ? \quad | \quad 8 \quad 7 \quad 5 \quad 9 \quad | \quad ? \quad ? \quad ?]^{\mathrm{T}}$$

$$\boldsymbol{v}_s' = [? \quad ? \quad | \quad 7 \quad 5 \quad 4 \quad 6 \quad | \quad ? \quad ? \quad ?]^{\mathrm{T}}$$

Podstawowe sposoby krzyżowania

- *Tworzenie potomka* $\mathbf{v}_r'$ *wg sekwencji genów chromosomu* $\mathbf{v}_s$ *począwszy od drugiego punktu krzyżowania*

$$\tilde{\boldsymbol{v}}_s = [8 \quad 2 \quad 3 \quad 1 \quad 9 \quad 7 \quad 5 \quad 4 \quad 6]^{\mathrm{T}}$$

- *Usunięcie z chromosomu* $\tilde{\mathbf{v}}_s$ *genów o allelach powtarzających się dla potomka* $\mathbf{v}_r'$

$$\tilde{\boldsymbol{v}}_s = [\quad 2 \quad 3 \quad 1 \qquad\qquad 4 \quad 6]^{\mathrm{T}}$$

Podstawowe sposoby krzyżowania

- *Uzyskany ciąg genów z* $\tilde{\boldsymbol{v}}_s$ *umieszczany w potomku* $\boldsymbol{v}_r^{'}$ *począwszy od pozycji określonej drugi punktem krzyżowania*

$$\boldsymbol{v}_r^{'} = [4 \quad 6 \quad | \quad 8 \quad 7 \quad 5 \quad 9 \quad | \quad 2 \quad 3 \quad 1]^{\mathrm{T}}$$

- *Potomek* $\boldsymbol{v}_s^{'}$ *określany podobnie*

$$\tilde{\boldsymbol{v}}_r = [1 \quad 3 \quad 6 \quad 4 \quad 2 \quad 8 \quad 7 \quad 5 \quad 9]^{\mathrm{T}}$$

$$\tilde{\boldsymbol{v}}_r = [1 \quad 3 \qquad\qquad 2 \quad 8 \qquad\qquad 9]^{\mathrm{T}}$$

$$\boldsymbol{v}_s^{'} = [8 \quad 9 \quad | \quad 7 \quad 5 \quad 4 \quad 6 \quad | \quad 1 \quad 3 \quad 2]^{\mathrm{T}}$$

Podstawowe sposoby krzyżowania

- **Krzyżowanie CX** w odróżnieniu od PMX i OX nie wykorzystuje punków krzyżowania do generowania potomków

- Operacja ta polega na odpowiednich kopiowaniu ciągów genów z wybranego rodzica według pozycji określonych drugi rodzicem

Podstawowe sposoby krzyżowania

Przykład 8. Krzyżowanie CX

$$\boldsymbol{v}_r = [4 \quad 2 \quad 8 \quad 7 \quad 5 \quad 9 \quad 1 \quad 3 \quad 6]^{\mathrm{T}}$$

$$\boldsymbol{v}_s = [1 \quad 9 \quad 7 \quad 5 \quad 4 \quad 6 \quad 8 \quad 2 \quad 3]^{\mathrm{T}}$$

- *Rozpoczęcie tworzenia* $\boldsymbol{v}_r'$ *potomka poprzez skopiowania pierwszego genu z rodzica* $\boldsymbol{v}_r$

$$\boldsymbol{v}_r' = [4 \quad ? \quad ? \quad ? \quad ? \quad ? \quad ? \quad ? \quad ?]^{\mathrm{T}}$$

Podstawowe sposoby krzyżowania

- *Pozycja skopiowanego genu wyznacza w rodzicu* $\mathbf{v}_s$ *następny gen jaki należy skopiować do potomka* $\mathbf{v}_r'$ *zgodnie z jego pozycją określoną w rodzicu* $\mathbf{v}_r$

$$\mathbf{v}_r' = [4 \quad ? \quad ? \quad ? \quad ? \quad ? \quad 1 \quad ? \quad ?]^{\mathrm{T}}$$

$$\mathbf{v}_r' = [4 \quad ? \quad 8 \quad ? \quad ? \quad ? \quad 1 \quad ? \quad ?]^{\mathrm{T}}$$

- *Powyższa reguła realizowana do momentu uzyskania tzw. zamkniętej sekwencji genów*

$$\mathbf{v}_r' = [4 \quad ? \quad 8 \quad 7 \quad 5 \quad ? \quad 1 \quad ? \quad ?]^{\mathrm{T}}$$

Podstawowe sposoby krzyżowania

- *Kopiowanie genów z rodzica* $\mathbf{v}_s$ *określonych pozycjami nieokreślonymi potomka* $\mathbf{v}_r'$

$$\mathbf{v}_r' = [4 \quad 9 \quad 8 \quad 7 \quad 5 \quad 6 \quad 1 \quad 2 \quad 3]^{\mathrm{T}}$$

- *Wyznaczenie drugiego potomka*

$$\mathbf{v}_s' = [1 \quad ? \quad ? \quad ? \quad ? \quad ? \quad ? \quad ? \quad ?]^{\mathrm{T}}$$
$$\mathbf{v}_s' = [1 \quad ? \quad ? \quad ? \quad 4 \quad ? \quad ? \quad ? \quad ?]^{\mathrm{T}}$$
$$\mathbf{v}_s' = [1 \quad ? \quad 7 \quad 5 \quad 4 \quad ? \quad 8 \quad ? \quad ?]^{\mathrm{T}}$$
$$\mathbf{v}_s' = [1 \quad 2 \quad 7 \quad 5 \quad 4 \quad 9 \quad 8 \quad 3 \quad 6]^{\mathrm{T}}$$

Mutacja

- **Mutacja** - przypadkowe ‘zaburzenia’ w nowo powstałych rozwiązaniach (potomkach)
- Ze względu na obiekt modyfikacji, mutacje można podzielić na dwie grupy:
  - **mutacje genotypowe** (genetyczne) oraz
  - **mutacje fenotypowe** (parametryczne)
- Mutacja - zmiana alleli genów z założonym prawdopodobieństwem p_m całej populacji osobników

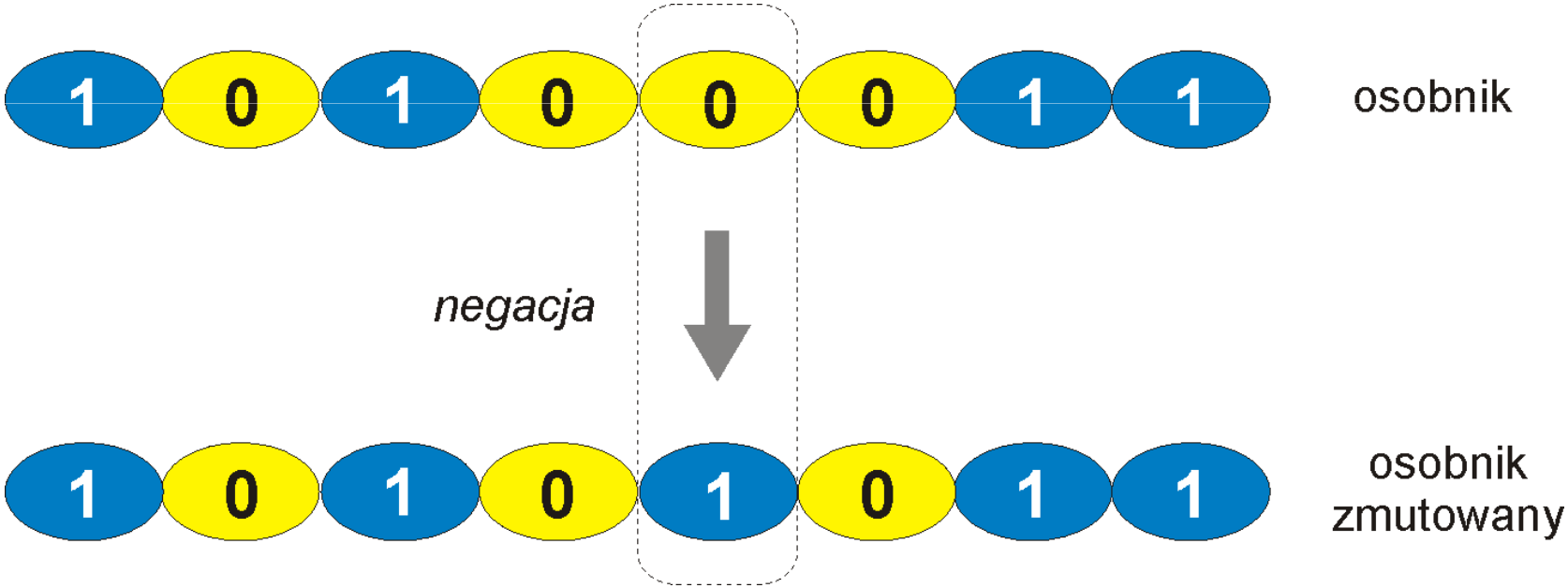
Mutacja genotypowa

- **Mutacja genotypowa** analogiczna do istniejącego mechanizmu tranzycji występującego przy mutacji nici DNA

- Dla kodowania binarnego - przypadkowa zmiana allelu genu w chromosomie na „przeciwny” (zanegowanie wybranego bitu kodu binarnego

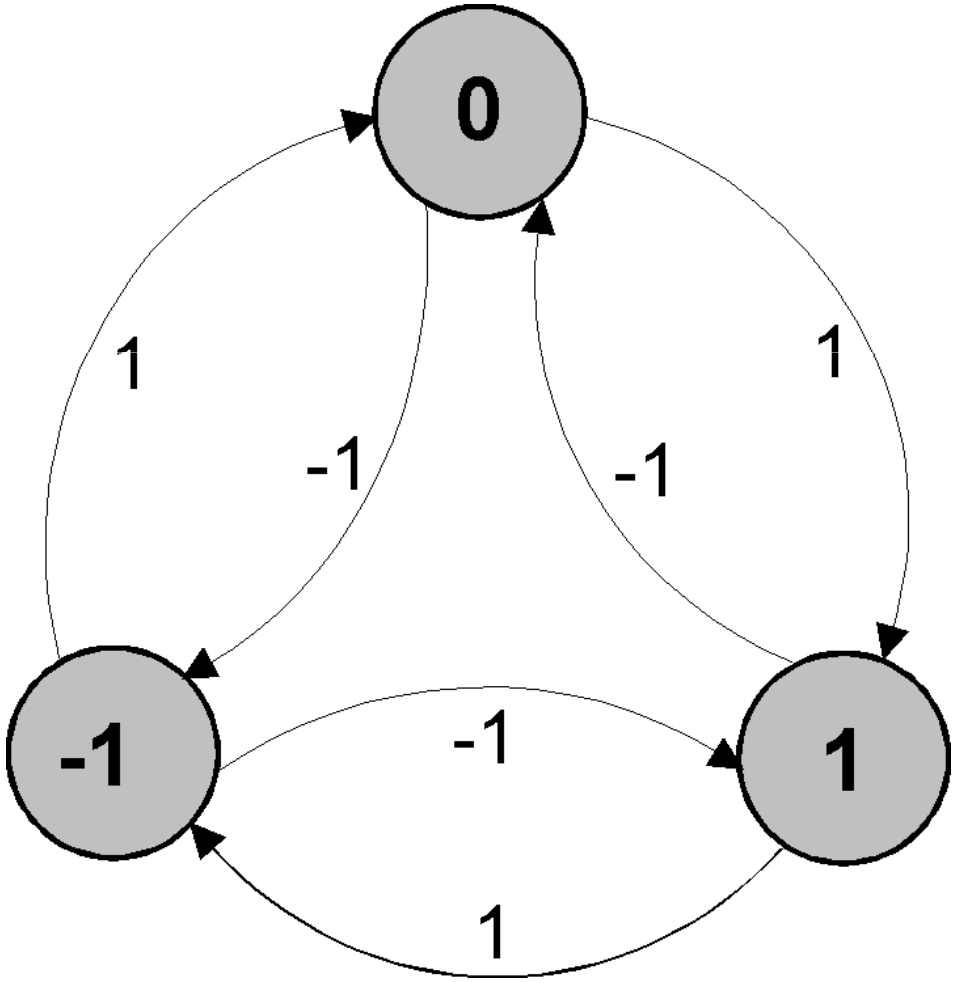
Mutacja genotypowa

Przykład mutacji genotypowej dla kodu binarnego



Mutacja genotypowa

Schemat mutacji genotypowej dla kodu trialleliczego



Mutacja fenotypowa

- **Mutacja fenotypowa** wprowadza zmianę wybranych parametrów
- Zmutowane osobniki są rozwiązaniami dopuszczalnymi (należą do dziedziny parametrów i liniowych ograniczeń na te parametry)
- Sposoby modyfikacji wartości parametru polegające na zmianie
  - **wartościowo-równomiernej**
  - **wartościowo-nierównomiernej**

Mutacja fenotypowa

- **Zmiana wartościowo-równomierna** - nadawanie wybranemu parametrowi v_{k_q} osobnika $\boldsymbol{v}_q$ losowej wartości z przedziału określającego dziedzinę mutowanego *k*-tego parametru z równomiernym rozkładem prawdopodobieństwa

- Potomek o nowym fenotypie:

$$\boldsymbol{v}_q^{'} = \left[v_1 \quad v_1 \quad \ldots \quad v_k^{'} \quad \ldots \quad v_n \right]^{\mathrm{T}}$$

Mutacja fenotypowa

- W zadaniach z liniowymi ograniczeniami dziedzina mutowanego parametru musi być określana wg narzuconych ograniczeń przy uwzględnieniu niezmienionych, pozostałych wartości parametrów w fenotypie

- Przymiotnik „równomierna” oznacza, że działanie mutacji jest niezależne od numeru iteracji algorytmu

Mutacja fenotypowa

- **Zmiana wartościowo-nierównomierna** – pozwala na dokładne dostrojenie parametrów

- Stochastyczne zaburzenie poprzednio ustalonej wartości wybranego parametru z zastosowaniem zawężającego się – wraz ze wzrostem numeru iteracji – dopuszczalnego przedziału

Mutacja fenotypowa

- Mutowany parametr v_{k_q} osobnika $\boldsymbol{v}_q$

$$v'_{k_q} = \begin{cases} v_{k_q} + \Delta(t, \bar{v}_k - v_{k_q}) & \text{(A)} \\ v_{k_q} - \Delta(t, v_{k_q} - \underline{v}_k) & \text{(B)} \end{cases}, \quad \Delta(t, \Delta v) = r \cdot \Delta v \left(1 - \frac{t}{t_{\max}}\right)^b$$

gdzie

$[\underline{v}_k, \bar{v}_k] \subset \Re$ - dziedzina parametru

t , $t_{\max}$ - kolejne, ostatnie pokolenie

b - współczynnik niejednorodności (b=2)

$r \in [0,1]$ - losowa z rozkładem równomiernym

Mutacja

- Inne mechanizmy mutacji stosowane dla genotypowej jak i fenotypowej struktury osobnika:
  - **inwersja** (ang. *inversion*)
  - **wstawianie** (ang. *insertion*)
  - **przenoszenie** (ang. *displacement*)
  - **wzajemna wymiana** (ang. *reciprocal exchange*)

Mutacja

- **Inwersja** - wylosowanie w chromosomie podciągu kodowego (sekwencji genów między dwoma losowo wybranymi pozycjami) i odwróceniu kolejności pozycji allei

- **Wstawianie** - wylosowanie pojedynczego genu i umiejscowieniu go w losowo wybranej pozycji w chromosomie osobnika

Mutacja

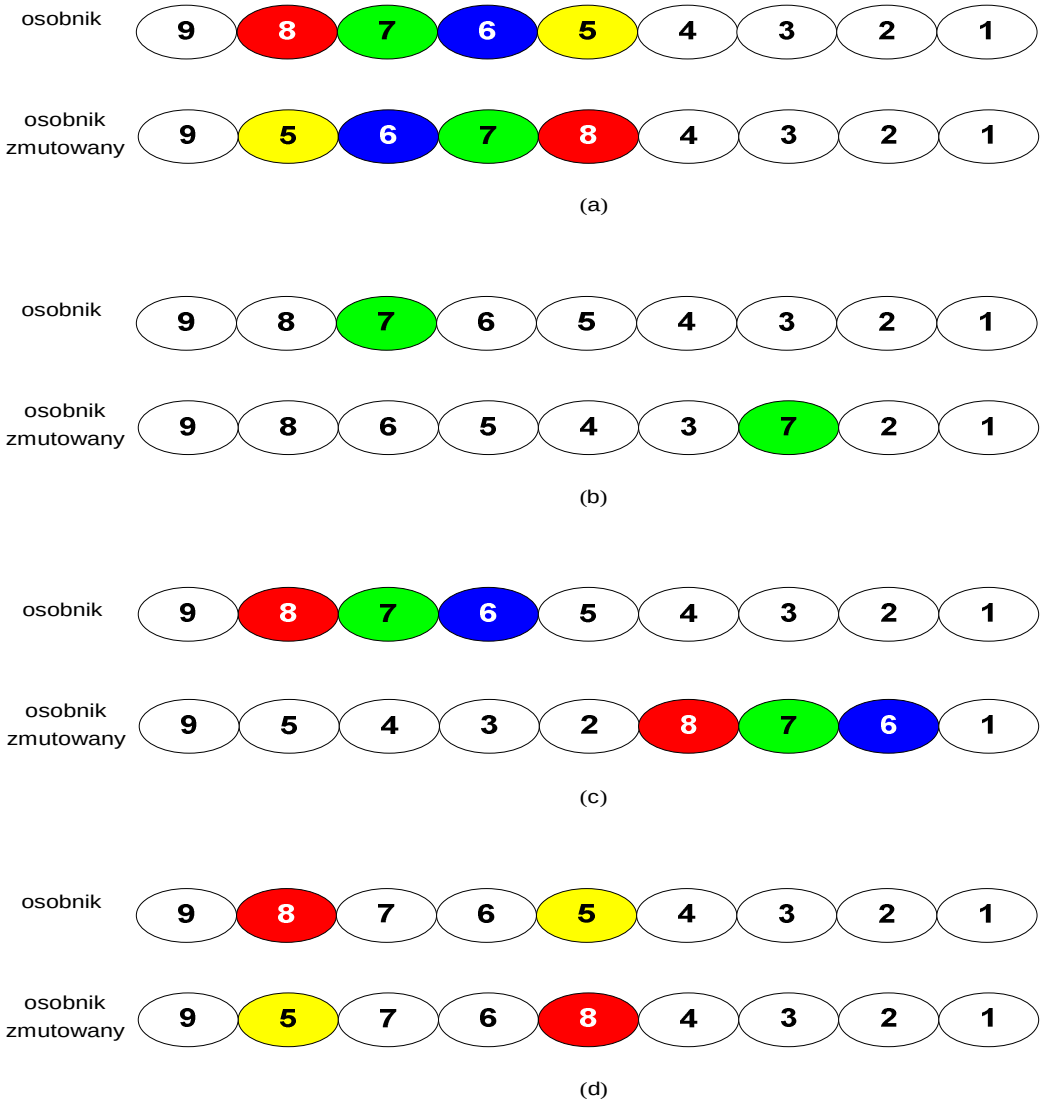
- **Przenoszenie** – wylosowywany podciąg kodowy wstawiany jest w losowo wybraną pozycje w chromosomie (uogólnienie mechanizmu wstawiania)

- **Wzajemna wymiana** – wylosowanie *k*-par genów do zamiany pozycji w chromosomie

Mutacja

Schematy operacji genetycznych:

a) inwersji
b) wstawiania
c) przenoszenia
d) wzajemnej wymiany



Strategie podstawień

- Proces zastępowania populacji uzyskanych w kolejnych pokoleniach

- Stosowane strategie podstawień
  - **reprodukcja pełna**
  - **reprodukcja częściowa**
  - **reprodukcja według czasu życia osobników**

Strategie podstawień

- **Strategia z reprodukcją pełną** - nowo powstała populacja o liczbie osobników *N* zastępuje całą poprzednią populację (najprostszy mechanizm)
- **Strategia z częściową reprodukcją -** zastąpienie wybranych osobników przez nowo powstałych potomków (możliwość posiadania wielopokoleniowego potomstwa)

Strategia z częściową reprodukcją

- Strategia z częściową reprodukcją realizowana na trzy sposoby poprzez:
  - **zastępowanie najgorszych osobników** (elityzm)
  - **zastępowanie osobników najbardziej ‘podobnych’ do potomnych** (metoda ze ściskiem)
  - **zastępowanie losowo wybranych osobników**

Strategia z częściową reprodukcją

- Usuwanie z populacji osobników z najniższym stopniem przystosowania i w ich miejsce wstawianie potomków lepiej przystosowanych
- Operacja ta polega na:
  - łącznym sortowaniu według niemalejących stopni przystosowania rodziców i potomków
  - wybraniu *N*-pierwszych najlepszych osobników
- Strategia elitarna może spowodować wystąpienie zjawiska przedwczesnej zbieżności (tworzenie elity)

Strategia z częściową reprodukcją

- **Metoda ze ściskiem** - usuwania osobników z największym podobieństwem w sensie:
  - genotypowym (w sensie odległości Hamminga)
  - fenotypowym (według normy euklidesowej)
- Potomek porównywany jest genotypowo lub fenotypowo z k-rodzicami wybieranymi losowo z rozkładem równomiernym ($k \leq N$)
- Najbardziej podobny rodzic usuwany z populacji i zastępowany potomkiem
- Wartość k nazywana współczynnikiem ścisku lub stłoczenia osobników w populacji o rozmiarze *N*

Strategia z reprodukcją wg czasu życia

- Przeprowadzana jest na podstawie przypisanego osobnikom dodatkowego wskaźnika zwanego czasem życia
- Wskaźnik życia zmniejsza się w kolejnych pokoleniach
- Z populacji usuwane są osobniki, których czas życia jest mniejszy lub równy zeru
- Nowa populacja jest ustalana na podstawie wyboru najlepszych (według stopnia przystosowania) osobników rodzicielskich i potomnych

Strategia z reprodukcją wg czasu życia

- Przypisanie osobnikom dodatkowego wskaźnika - **czasu życia**
- Z populacji usuwane są osobniki o czasie życia mniejszym lub równy zeru
- Nowa populacja ustalana na podstawie wyboru najlepszych rodziców i potomków
- Sposoby strategii wg:
  - **stałego czasu życia osobnika**
  - **zmiennego czasu życia**

Strategia z reprodukcją wg czasu życia

- **Reprodukcja według stałego czasu życia osobników -** przypisanie wszystkim osobnikom pewnego arbitralnie dobranego maksymalnego czasu życia
- **Reprodukcja według zmiennego czasu życia osobników**:
  - populacje w kolejnych pokoleniach mogą przyjmować różne rozmiary
  - pula rodzicielska jest wyłaniania poprzez losowy wybór osobników z populacji z rozkładem równomiernym

Strategia z reprodukcją wg czasu życia

- Czas życia osobnika $l(x_i)$ określany na podstawie jego stopnia przystosowania za pomocą trzech możliwych przekształceń:
  - **liniowego**
  - **proporcjonalnego**
  - **biliniowego**

Strategia z reprodukcją wg czasu życia

- Stosowane oznaczenia:

l_{min} - minimalny czas życia

l_{max} - maksymalny czas życia

$l_{avg} = \frac{l_{max} + l_{min}}{2}$ - średni czas życia

$$f_{min}(\Delta t) = \min_{t=t_1,\ldots,t_2} \{ \min_{i=1,2,\ldots,N} f(\boldsymbol{x}_i) \}$$

$$f_{avg}(\Delta t) = \frac{1}{\Delta t} \sum_{t=t_1}^{t_2} \left(\frac{1}{N} \sum_{i=1}^{N} f(\boldsymbol{x}_i) \right)$$

$$f_{max}(\Delta t) = \max_{t=t_1,\ldots,t_2} \{ \max_{i=1,2,\ldots,N} f(\boldsymbol{x}_i) \}$$

Strategia z reprodukcją wg czasu życia

- Stosowane oznaczenia:

  t_1, t_2 - numery pokoleń wyznaczające okno czasowe $\Delta t = t_2 - t_1$
- Wartość Δt wpływa na nacisk selektywny
- Małe okno czasowe powoduje większy nacisk selektywny
- Duże okno czasowe wartość daje mniejszy nacisk selektywny

Strategia z reprodukcją wg czasu życia

- Przekształcenie za pomocą funkcji:
  - liniowej

$$l(\boldsymbol{x}_i) = l_{min} + \frac{l_{max} - l_{min}}{f_{max}(\Delta t) - f_{min}(\Delta t)} f(\boldsymbol{x}_i)$$

  - proporcjonalnej

$$l(\boldsymbol{x}_i) = \min\left\{ l_{max},\ l_{min} + \frac{l_{avg} - l_{min}}{f_{avg}(\Delta t) - f_{min}(\Delta t)} f(\boldsymbol{x}_i) \right\}$$

Strategia z reprodukcją wg czasu życia

- Przekształcenie za pomocą funkcji:
  - biliniowej

$$l(\mathbf{x}_i) = \begin{cases} l_{min} + \dfrac{l_{avg} - l_{min}}{f_{avg}(\Delta t) - f_{min}(\Delta t)} f(\mathbf{x}_i), & f(\mathbf{x}_i) < f_{avg}(\Delta t) \\ l_{avg} + \dfrac{l_{max} - l_{avg}}{f_{max}(\Delta t) - f_{avg}(\Delta t)} f(\mathbf{x}_i), & f(\mathbf{x}_i) \geq f_{avg}(\Delta t) \end{cases}$$

Strategia z reprodukcją wg czasu życia

zmiana \ funkcja	liniowa	proporcjonalna	biliniowa
$\downarrow f_{min}(\Delta t)$	zwiększenie średniego czasu życia oraz osłabienie nacisku selektywnego	zwiększenie średniego czasu życia oraz osłabienie nacisku selektywnego	osłabienie nacisku selektywnego dla osobników o stopniu przystosowania poniżej średniej
$\downarrow f_{avg}(\Delta t)$	brak wpływu	zwiększenie liczby osobników o maksymalnym czasie życia	zmniejszenie nacisku selektywnego dla osobników lepszych oraz zwiększenie nacisku selektywnego dla osobników gorszych
$\uparrow f_{avg}(\Delta t)$	brak wpływu	zmniejszenie liczby osobników o maksymalnym czasie życia	zwiększenie nacisku selektywnego dla osobników lepszych oraz zmniejszenie nacisku selektywnego dla osobników gorszych
$\uparrow f_{max}(\Delta t)$	zmniejszenie średniego czasu życia oraz zmniejszenie nacisku selektywnego	brak wpływu	zmniejszenie nacisku selektywnego dla osobników powyżej przeciętnej

Skalowanie przystosowania

- Zjawisko przedwczesnej zbieżności (ang. *premature convergance*) do suboptymalnego rozwiązania
- Metody skalowania przystosowania:
  - liniowe
  - z odchyleniem standardowym (sigma-obcinające)
  - wykładnicze (potęgowe)

Skalowanie liniowe

- Zmodyfikowane przystosowanie osobników

$$f_s(\boldsymbol{x}_i) = a \cdot f(\boldsymbol{x}_i) + b$$

- Warunki dobierania parametrów *a* i *b*:
  - średnie przystosowanie populacji jest niezmienne po skalowaniu
  - zmodyfikowane maksymalne przystosowania osobników równe wielokrotności średniego przystosowania populacji
  - przeskalowane wartości funkcji przystosowania nie mogą przyjmować wartości ujemnych.

Schemat skalowania liniowego

1. Znajdź w populacji osobnika o maksymalnym i minimalnym przystosowaniu

$$f_{max} = \max_{i \in \{1,2,\ldots,N\}} f(\boldsymbol{x}_i)$$

$$f_{min} = \min_{i \in \{1,2,\ldots,N\}} f(\boldsymbol{x}_i)$$

2. Wyznacz średnie przystosowanie populacji

$$f_{avg} = \frac{1}{N} \sum_{i=1}^{N} f(\boldsymbol{x}_i)$$

Schemat skalowania liniowego

3. Przeskaluj stopnie przystosowania osobników

$$f_s(\boldsymbol{x}_i) = a \cdot f(\boldsymbol{x}_i) + b$$

a i *b* współczynniki skalowania linowego wartości wyznaczane następująco:

– sprawdź warunek skalowania

$$f_{min} > \frac{c \cdot f_{avg} - f_{max}}{c - 1} \qquad (*)$$

c - współczynnik zwielokrotnienia (liczba kopii)

Schemat skalowania liniowego

– jeżeli warunek skalowania (*) spełniony to

$$a = \frac{f_{avg}(c-1)}{f_{max} - f_{avg}}, \quad b = -f_{avg}(a+1)$$

w przeciwnym przypadku

$$a = \frac{f_{avg}}{f_{avg} - f_{min}}, \quad b = -af_{min}$$

- W których pokoleniach cyklu GA są wybierane odpowiednie parametry *a* i *b?*

Skalowanie σ-obcinające

- Zmodyfikowane przystosowanie osobników

$$f_s(\mathbf{x}_i) = f(\mathbf{x}_i) + (f_{avg} - c \cdot \sigma)$$

gdzie

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (f(\mathbf{x}_i) - f_{avg})^2}$$

c - liczba naturalna (zwykle od 1 do 5)

Skalowanie wykładnicze

- Zmodyfikowane przystosowanie osobników

$$f_s(\mathbf{x}_i) = f(\mathbf{x}_i)^k$$

gdzie

k - liczba zwykle bliską jedności $(k = 1.005)$

- W zależności od problemu optymalizacji wartość *k* powinna być dobrana odpowiednio
- Należy zadbać by zmieniała się podczas cyklu ewolucyjnego

Niszowanie

- Zachowanie gorzej przystosowanych osobników by populacja była różnorodna
- Pozwala utrzymać stałą liczbę istniejących gatunków zarówno:
  - bardziej licznych (lepiej przystosowanych)
  - mniej licznych (słabiej przystosowanych).
- Nisza w GA jest pewnym obszarem, w którym znajdują się osobniki bliskie parametrycznie (fenotypowo, geometrycznie), które powinny mieć także podobne stopnie przystosowania

Mechanizm niszowania

- Funkcja bliskości (ang. *sharing function*)

$$\delta_{ij} = \begin{cases} 1 - \left\| \boldsymbol{x}_i - \boldsymbol{x}_j \right\|_P & \text{jeśli} \quad 0 \le \left\| \boldsymbol{x}_i - \boldsymbol{x}_j \right\|_P < 1 \\ 0 & \text{jeśli} \quad \left\| \boldsymbol{x}_i - \boldsymbol{x}_j \right\|_P \ge 1 \end{cases}$$

gdzie

$$\left\| \boldsymbol{x}_i - \boldsymbol{x}_j \right\|_P = \sqrt{(\boldsymbol{x}_i - \boldsymbol{x}_j)^{\mathrm{T}} \boldsymbol{P}^{-2} (\boldsymbol{x}_i - \boldsymbol{x}_j)}$$

$$\boldsymbol{P} = diag\{\phi_1 \quad \phi_2 \quad \dots \quad \phi_n\}$$

$\phi_k \ (k = 1,2,\dots,n)$ - *k*-ta średnica hiperelipsoidy o środku w *i*-tym osobniku

Mechanizm niszowania

- *k*-ta średnica hiperelipsoidy

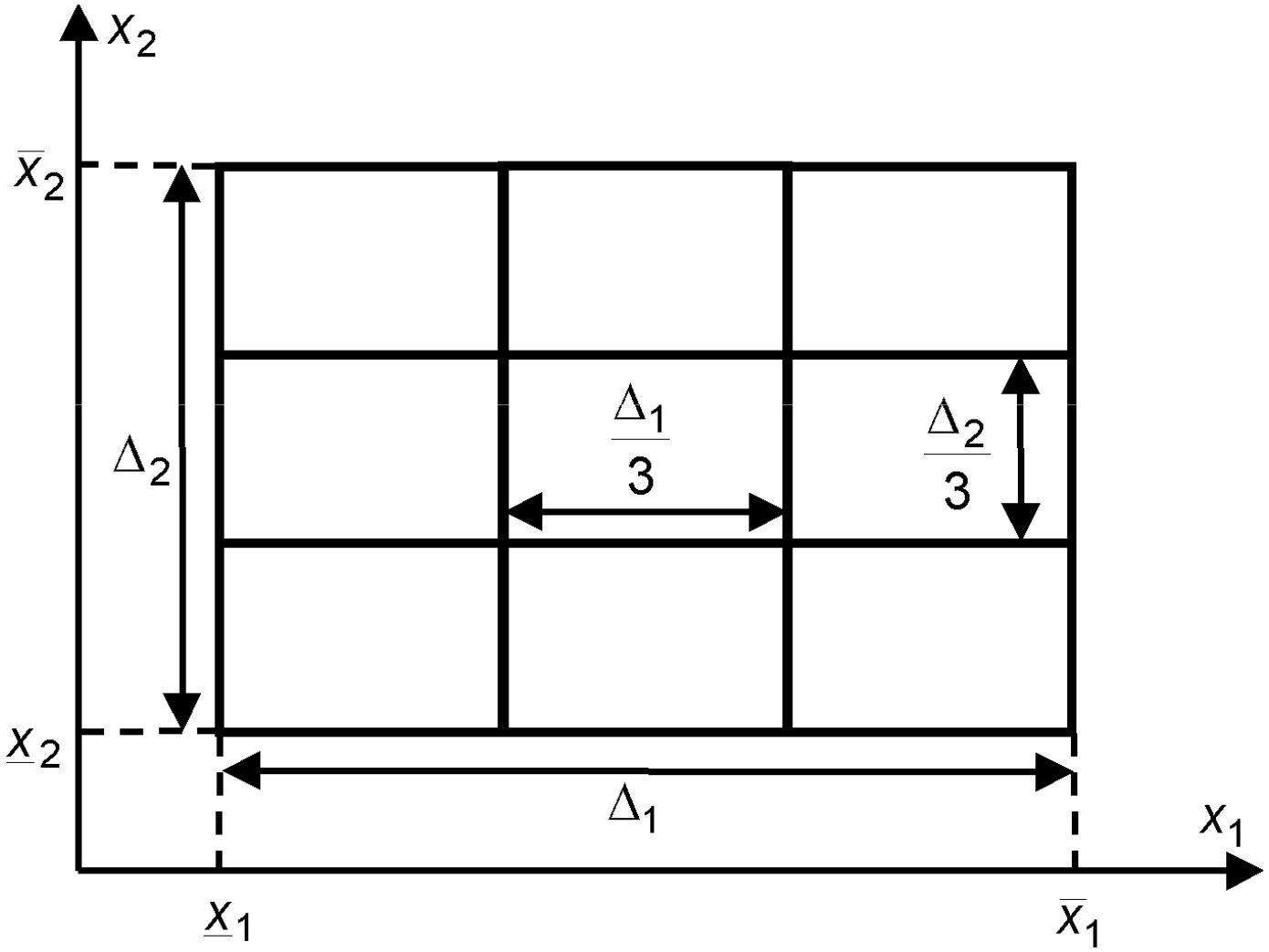
$$\phi_k = \frac{\Delta_k}{\varepsilon}, \quad k = 1,2,\ldots,n$$

gdzie

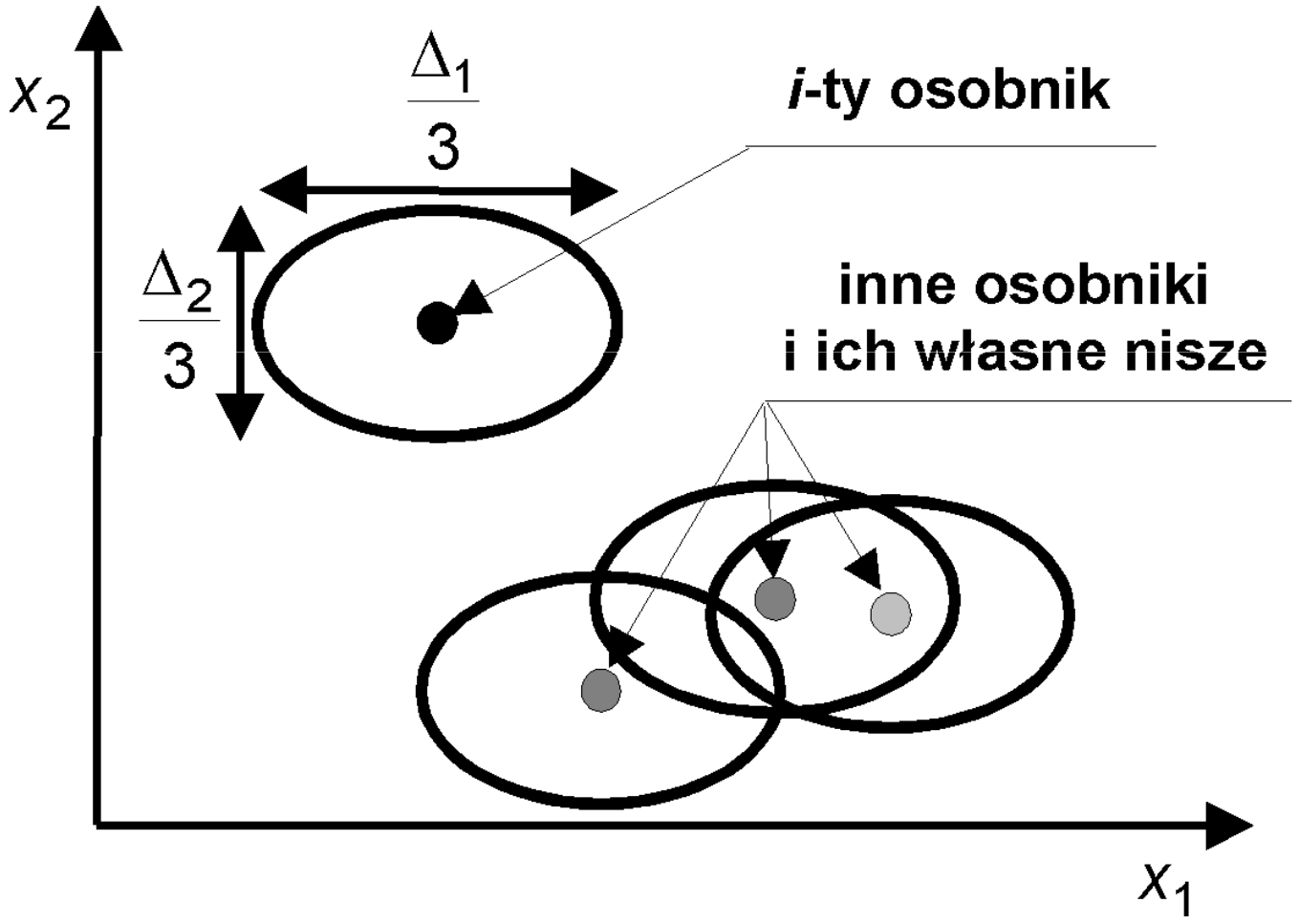
Δ_k - niezerowym zakresem poszukiwań *k*-tego parametru

ε - liczba części, na jaką dzieli się zakres Δ_k

Mechanizm niszowania



Mechanizm niszowania



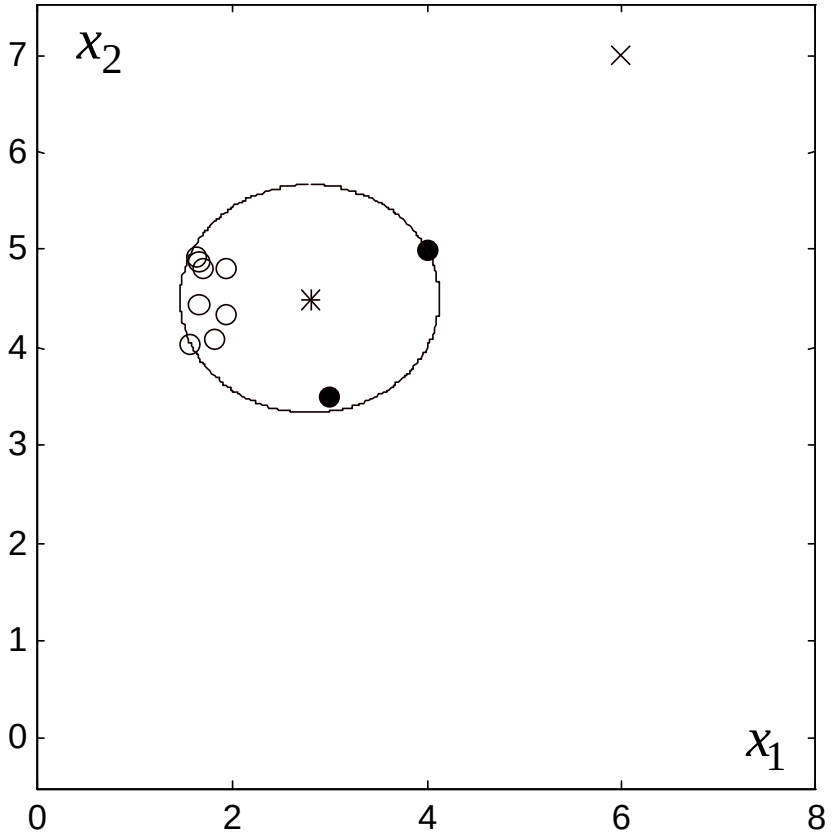
Mechanizm niszowania

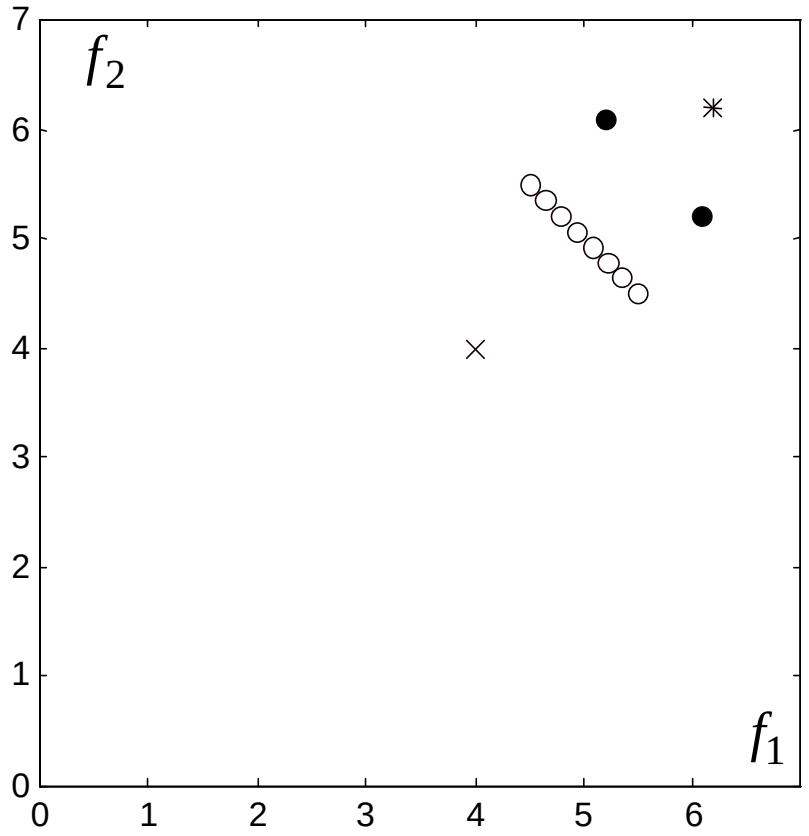
- Modyfikacja stopnia przystosowania lub rangi każdego osobnika znajdującego się we ‘własnej’ niszy według następującego przepisu

$$\tilde{f}(x_i) = \frac{f(x_i)}{\sum_{j=1}^{N} \delta_{ij}} \qquad \tilde{r}(x_i) = \frac{r(x_i)}{\sum_{j=1}^{N} \delta_{ij}}$$

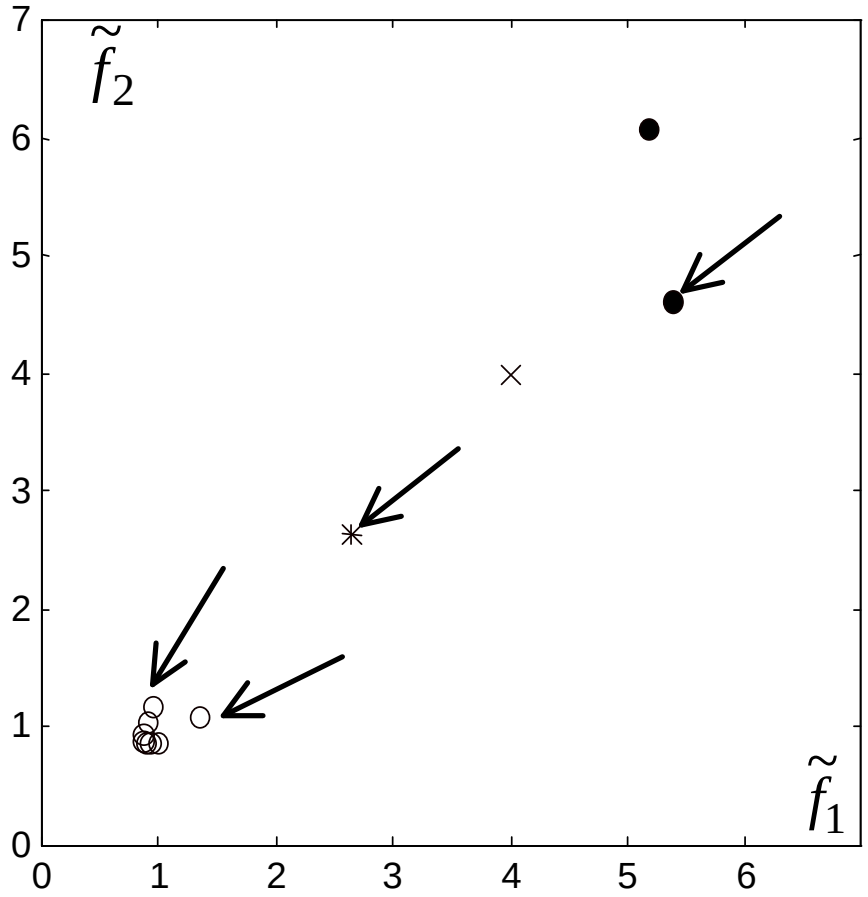
- Suma w obu mianownikach wyraża efektywną liczbę osobników w niszy własnej o środku opisanym przez fenotyp rozważanego *i*-tego osobnika

Mechanizm niszowania





Mechanizm niszowania



Teoria schematów

- Metodologię procesu poszukiwań GA można interpretować w kategoriach schematów
- Schemat - wzorzec opisujący podzbiór sekwencji kodowych podobnych pod względem pozycji o ustalonych wartościach zwanych pozycjami ustalonymi
- Schemat składa się z cyfr kodu binarnego i dodatkowego znak “#”
- Symboli “#” oznacza nieustaloną wartość pozycji

Teoria schematów

Przykład 9. Schematy

$$S_1 = [0 \quad 0 \quad 1 \quad \# \quad 1 \quad 1 \quad 0 \quad \# \quad 0 \quad 1 \quad 1]^{\mathrm{T}}$$

Do schematu S_1 pasują cztery chromosomy

$$\boldsymbol{v}_1 = [0 \quad 0 \quad 1 \quad 0 \quad 1 \quad 1 \quad 0 \quad 0 \quad 0 \quad 1 \quad 1]^{\mathrm{T}}$$

$$\boldsymbol{v}_2 = [0 \quad 0 \quad 1 \quad 0 \quad 1 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 1]^{\mathrm{T}}$$

$$\boldsymbol{v}_3 = [0 \quad 0 \quad 1 \quad 1 \quad 1 \quad 1 \quad 0 \quad 0 \quad 0 \quad 1 \quad 1]^{\mathrm{T}}$$

$$\boldsymbol{v}_4 = [0 \quad 0 \quad 1 \quad 1 \quad 1 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 1]^{\mathrm{T}}$$

Teoria schematów

- Schemat posiada następujące dwie istotne własności:
  - **rząd**
  - **rozpiętość**
- **Rzędem schematu** nazywa się liczbę ustalonych pozycji, tzn. liczbę 0 i 1 w schemacie
- **Rozpiętością schematu** nazywa się odległość między skrajnymi ustalonymi pozycjami, tzn. genami o ustalonej wartości (allelu) 0 lub 1

Teoria schematów

- Liczba osobników pasujących do danego schematu *S* po selekcji metodą proporcjonalną

$$\xi_{t+1}(S) = \xi_t(S) \cdot N \cdot \frac{F_t(S)}{\sum_{i=1}^{N} f_t(\boldsymbol{x}_i)}$$

gdzie

$$F_t(S) = \frac{1}{j} \sum_{i=1}^{j} f_t(\boldsymbol{x}_i)$$

Teoria schematów - efekt selekcji

- Średni stopień przystosowania populacji

$$f_{avg}(t) = \frac{1}{N} \sum_{i=1}^{N} f_t(\boldsymbol{x}_i)$$

- Zatem

$$\xi_{t+1}(S) = \xi_t(S) \cdot \frac{F_t(S)}{f_{avg}(t)}$$

- Schematy o przystosowaniu wyższym niż średnie przystosowanie populacji mają większą liczbę reprezentantów w następnym pokoleniu

Teoria schematów - efekt selekcji

- Jeżeli schemat *S* ma przystosowanie większe od średniego przystosowania populacji o pewną stałą ε

$$F_t(S) = f_{avg}(t) + \varepsilon \cdot f_{avg}(t)$$

$$\xi_{t+1}(S) = (1+\varepsilon) \cdot \xi_t(S)$$

- Poczynając od 0 i zakładając, że ε nie zmienia się

$$\xi_{t+1}(S) = \xi_0(S) \cdot (1+\varepsilon)^t$$

Teoria schematów - efekt selekcji

Przykład 9. Schematy i ich przystosowanie przed selekcją metodą proporcjonalną

Lp.	S_i	$\boldsymbol{v_i}$	$F_t(S_i)$	$\xi_t(S_i)$
1	[#### 0## #####]^T	$\boldsymbol{v_3}$=[0111 011 10010]^T $\boldsymbol{v_4}$=[0100 010 01000]^T	1067	1.380
2	[0### ### 0####]^T	$\boldsymbol{v_4}$=[0100 010 01000]^T	1216	1.573
3	[#### #10 #####]^T	$\boldsymbol{v_1}$=[0110 110 11110]^T $\boldsymbol{v_4}$=[0100 010 01000]^T	714.5	0.959

Teoria schematów - efekt selekcji

Schematy po selekcji metodą proporcjonalną

Lp.	S_i	$\boldsymbol{v}_i$	$F_{t+1}(S_i)$	$\xi_{t+1}(S_i)$
1	[#### 0## #####]^T	$\boldsymbol{v}_3$=[0111 011 10010]^T $\boldsymbol{v}_4$=[0100 010 01000]^T	1067	1.906
2	[0### ### 0####]^T	$\boldsymbol{v}_4$=[0100 010 01000]^T	1216	2.474
3	[#### #10 #####]^T	$\boldsymbol{v}_4$=[0100 010 01000]^T	1216	2.474

Teoria schematów - efekt krzyżowania

- Zakłada się krzyżowanie jednopunktowe dla osobników hapliodalnych kodowanych binarnie
- Prawdopodobieństwo destrukcji (utraty) $p_d(S)$ oraz przeżycia (zachowania) $\bar{p}_d(S)$ schematu *S*

$$p_d(S) = \frac{\delta(S)}{m-1} \qquad \bar{p}_d(S) = 1 - \frac{\delta(S)}{m-1}$$

gdzie

$\delta(S)$ - rozpiętość schematu *S*

m - długość chromosomu pasującego do *S*

Teoria schematów - efekt krzyżowania

Przykład 10. Krzyżowanie jednopunktowe pary rodzicielskiej dla wybranego punktu krzyżowania między bitem 5 a 6

$$\boldsymbol{v}_1 = [0 \quad 0 \quad 1 \quad 0 \quad 1 \quad 1 \quad | \quad 0 \quad 0 \quad 0 \quad 1 \quad 1]^{\mathrm{T}}$$

$$\boldsymbol{v}_2 = [1 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \quad | \quad 1 \quad 0 \quad 0 \quad 1 \quad 0]^{\mathrm{T}}$$

Rozważmy dwa schematy

$$S_1 = [\# \quad \# \quad 1 \quad \# \quad \# \quad \# \quad | \quad \# \quad \# \quad \# \quad 1 \quad 1]^{\mathrm{T}}$$

$$S_2 = [\# \quad \# \quad \# \quad 0 \quad 1 \quad \# \quad | \quad \# \quad \# \quad \# \quad \# \quad \#]^{\mathrm{T}}$$

Teoria schematów - efekt krzyżowania

Schematy charakteryzują się następującymi rzędami i rozpiętościami:

$$\rho(S_1) = 3, \quad \delta(S_1) = 8$$

$$\rho(S_1) = 2, \quad \delta(S_1) = 1$$

Potomkowie po krzyżowaniu jednopunktowym

$$\boldsymbol{v}_1' = [0 \quad 0 \quad 1 \quad 0 \quad 1 \quad 1 \quad | \quad 1 \quad 0 \quad 0 \quad 1 \quad 0]^{\mathrm{T}}$$

$$\boldsymbol{v}_2' = [1 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \quad | \quad 0 \quad 0 \quad 0 \quad 1 \quad 1]^{\mathrm{T}}$$

Teoria schematów - efekt krzyżowania

Prawdopodobieństwa destrukcji i przeżycia schematów S_1 i S_2 wynoszą odpowiednio

$$p_d(S_1) = \frac{8}{10}, \quad \bar{p}_d(S_1) = \frac{2}{10}$$

$$p_d(S_2) = \frac{1}{10}, \quad \bar{p}_d(S_2) = \frac{9}{10}$$

Wniosek: Większa szansa przeżycia schematów o małej rozpiętości

Teoria schematów - efekt krzyżowania

- Jeśli operacja krzyżowania jest wykonywana z prawdopodobieństwem krzyżowania $0<p_c<1$
- Wówczas prawdopodobieństwo zniszczenia oraz zachowania schematu *S* ma postać

$$p_d(S) \geq p_c \cdot \frac{\delta(S)}{m-1}$$

$$\overline{p}_d(S) \geq p_c \cdot \frac{\delta(S)}{m-1}$$

Teoria schematów - efekt mutacji

- Schemat *S* przeżywa mutację jeśli ustalone pozycje (geny o wartościach 0 lub 1) zostają zachowane
- Gen mutowany z prawdopodobieństwem mutacji p_m
- Prawdopodobieństwo przeżycia schematu *S*

$$p_{s_m}(S) = (1 - p_m)^{\rho(S)}$$

gdzie

$\rho(S)$ - rząd schematu *S*

Teoria schematów - efekt mutacji

- Aproksymacja prawdopodobieństwo przeżycia schematu *S* (dla małych wartości $p_m << 1$)

$$p_{s_m}(S) \approx 1 - \rho(S) \cdot p_m$$

- Prawdopodobieństwa przeżycia z przykładu 10

$$p_{s_m}(S_1) \approx 1 - 3 \cdot p_m$$
$$p_{s_m}(S_2) \approx 1 - 2 \cdot p_m$$

- Wniosek: Schematy o małym rzędzie przetrwają mutację

Teoria schematów - twierdzenie

- Twierdzenie schematów

$$\xi_{t+1}(S) \geq \xi_t(S) \cdot \frac{F_t(S)}{f_{avg}(t)} \cdot \left(1 - p_c \cdot \frac{\delta(S)}{m-1} - \rho(S) \cdot p_m\right)$$

- Schematy o:
  - małej rozpiętości
  - niskiego rzędu
  - z wyższym stopniem przystosowania

  uzyskują wykładniczo rosnącą liczbę reprezentantów w następnych pokoleniach

Teoria schematów - twierdzenie

- **Hipoteza**: AG poszukuje rozwiązania optymalnego poprzez „zachowywanie” krótkich schematów niskiego rzędu o stopniu przystosowania wyższym od średniego
- Schematy takie nazwane są **blokami budującymi**
- Dowód oparty głównie na wynikach empirycznych uzyskanych dla szerokiego spektrum zastosowań GA

Teoria schematów - twierdzenie

- Przetwarzanie w toku genetycznej ewolucji co najmniej 2^m i $N \cdot 2^m$ co najwyżej schematów

  (populacja o *N* osobnikach kodowanych binarnie o długości chromosomu *m*)

- Nie wszystkie schematy mają szansę na przetrwanie:
  - operacja krzyżowania niszczy schematy o dużej rozpiętości
  - mutacja niszczy schematy o wyższym rzędzie

Teoria schematów - twierdzenie

- Przy względnie długich chromosomach ($2^m > N^2$) efektywna liczba schematów przetwarzana w każdym pokoleniu przez AG wynosi N^3
- Własność AG:
  - ukryta równoległość (ang. *implicit parallelism*)
- AG działając na zbiorze N osobników dokonuje przetwarzania schematów N^3

Symulacja GA

- Maksymalizacja odwróconej funkcji Shekhel'a - 'wilcze doły'

$$\max_{x_1,x_2} g(x_1,x_2) = -\left[\frac{1}{500} + \sum_{j=1}^{25} \frac{1}{j + (x_1 - a_{ij})^6 + (x_2 - a_{ij})^6}\right]^{-1}$$

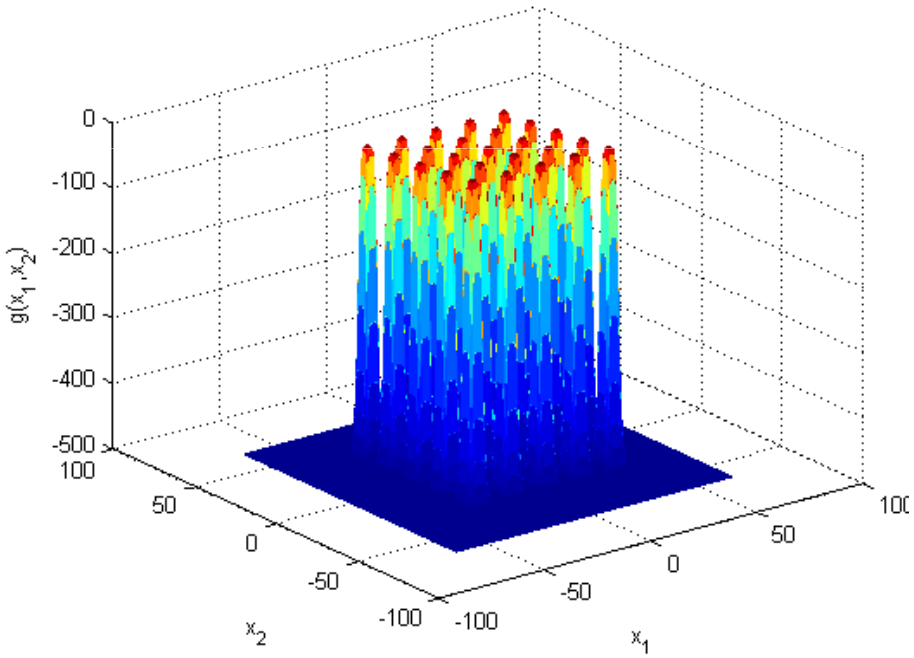
$$\boldsymbol{A} = \begin{bmatrix} -32 & -16 & 0 & 16 & 32 & -32 & \dots & 0 & 16 & 32 \\ -32 & -32 & -32 & -32 & -32 & -16 & \dots & 32 & 32 & 32 \end{bmatrix}$$

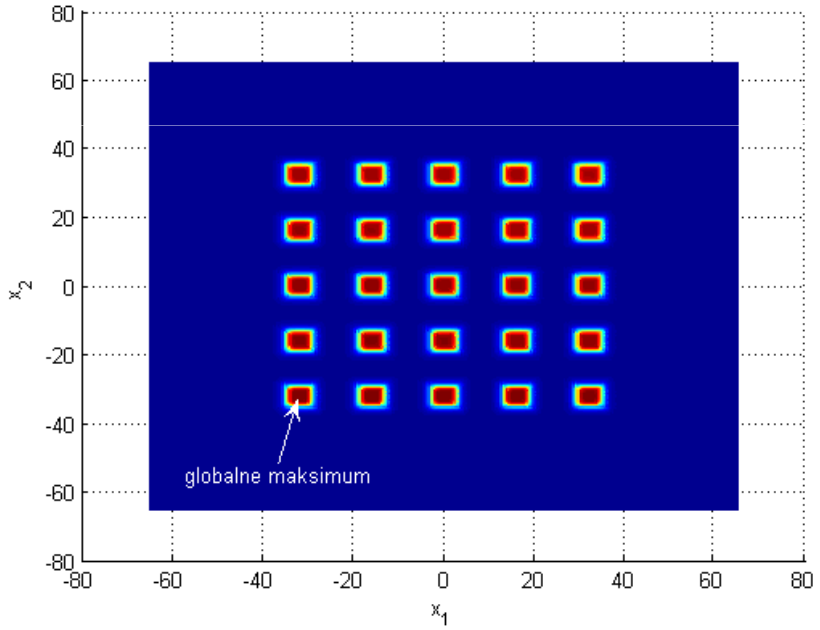
Symulacja GA

- *W zastosowanym algorytmie genetycznym założono następujące parametry:*
  - *dziedzina poszukiwań* $x_{1,2} \in [-65, 65]$
  - *50 osobników w populacji*
  - *24 bity na każdy parametr*
  - *dwupunktowe krzyżowanie z* $p_c = 0.8$
  - *binarna mutacja z* $p_m = 0.01$
  - *liczba pokoleń równa 50*

Symulacja GA

Odwrócona funkcja Shekhel'a - 'wilcze doły'

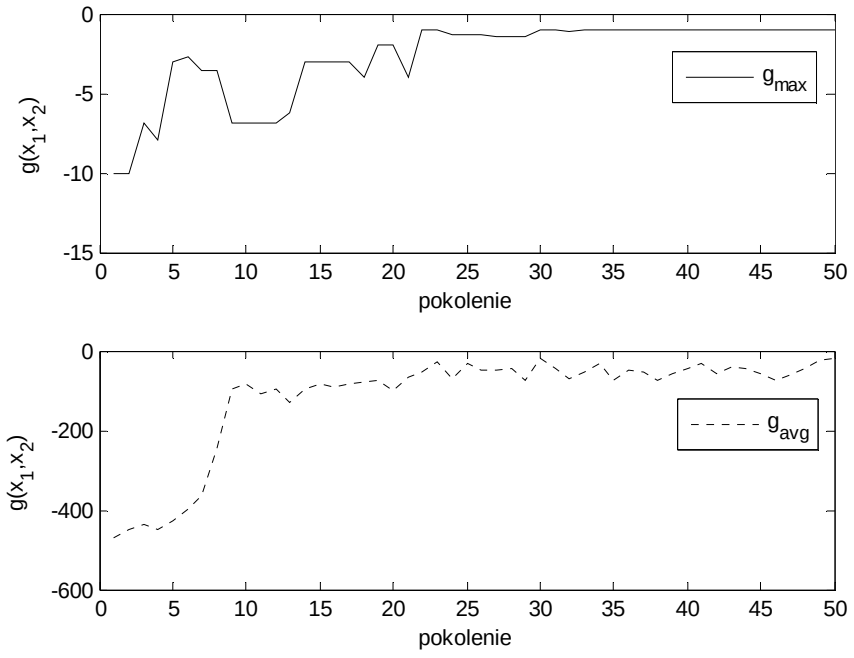




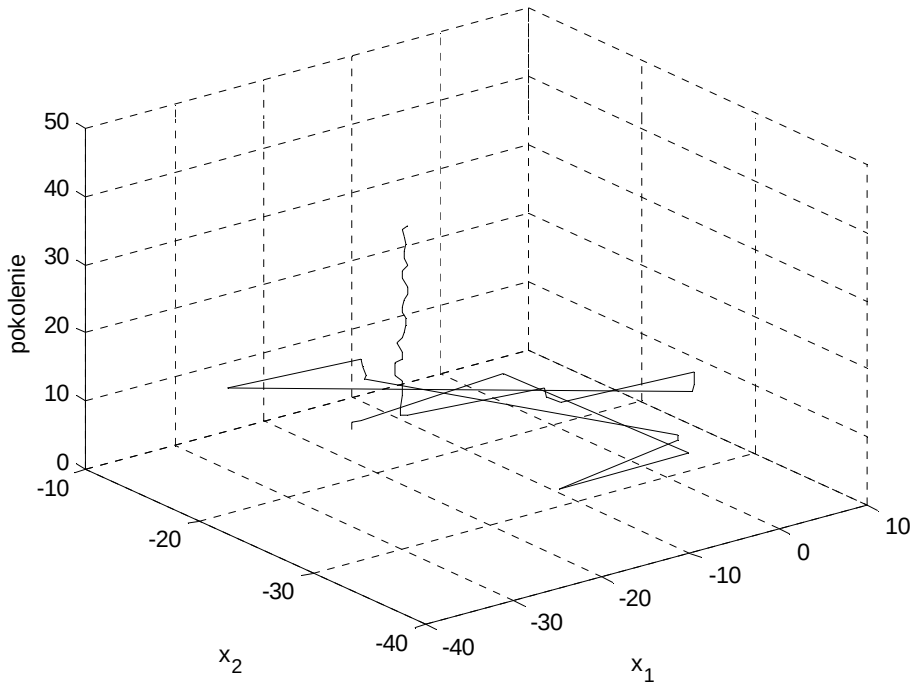
Symulacja GA

Przebiegi algorytmu genetycznego dla:

a) maksymalnej i średniej wartości funkcji zysku
b) poszukiwanych parametrów.



(a)



(b)

Symulacja GA

- Maksymalizacja odwróconej funkcji skokowej

$$\max_{x_1,x_2} \; g(x_1,x_2) = (\lfloor x_1 + 0.5 \rfloor)^2 + (\lfloor x_2 + 0.5 \rfloor)^2$$

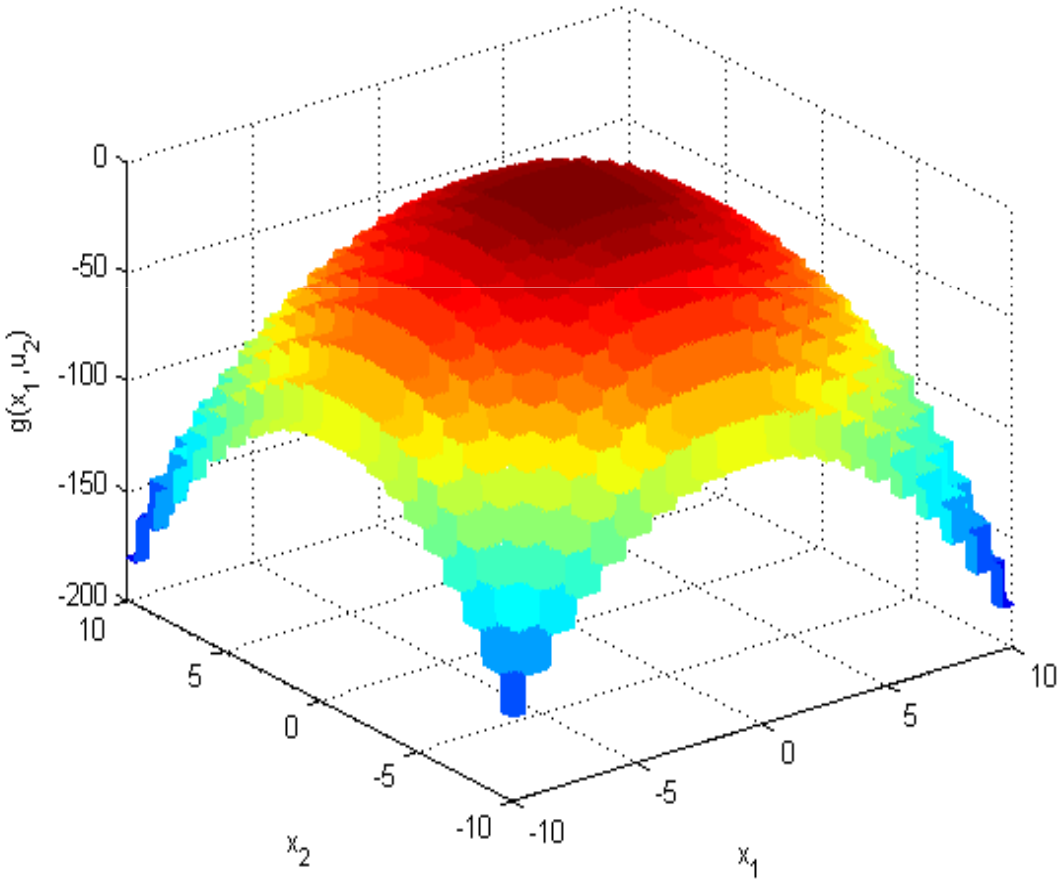
- Globalne maksimum w punkcie $\begin{bmatrix} x_1^* & x_2^* \end{bmatrix} = [0 \quad 0]$ o wartości funkcji $g(0,0) = 0$

Symulacja GA

- *W zastosowanym algorytmie genetycznym założono następujące parametry:*
  - *dziedzina poszukiwań* $x_{1,2} \in [-100, 100]$
  - *60 osobników w populacji*
  - *24 bity na każdy parametr*
  - *dwupunktowe krzyżowanie z* $p_c = 0.8$
  - *binarna mutacja z* $p_m = 0.01$
  - *liczba pokoleń równa 50*

Symulacja GA

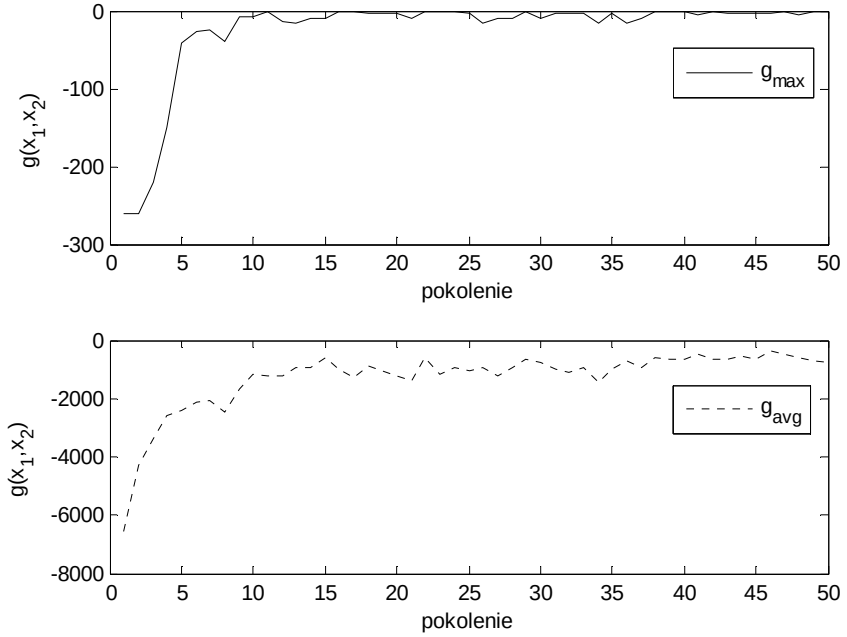
Odwrócona funkcja skokowa



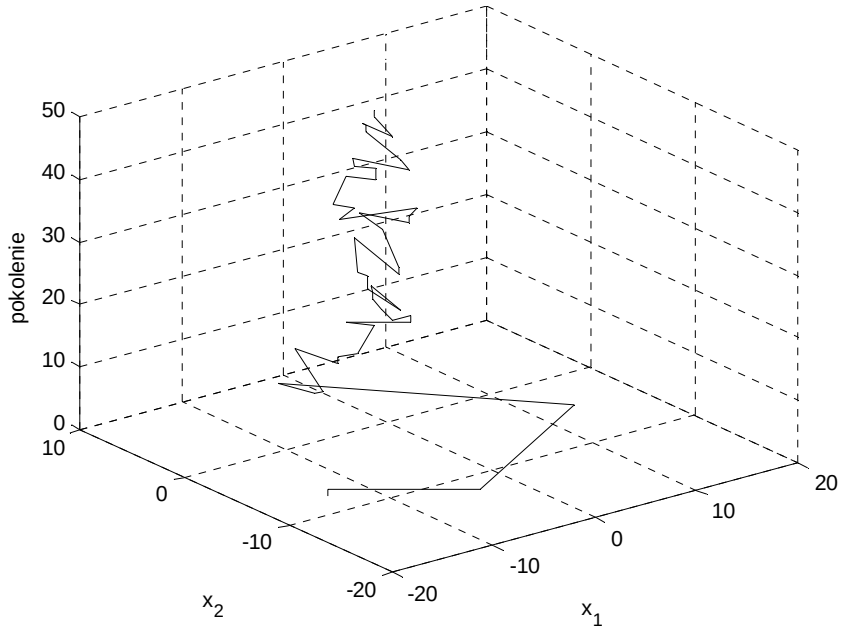
Symulacja GA

Przebiegi algorytmu genetycznego dla:

a) maksymalnej i średniej wartości funkcji zysku
b) poszukiwanych parametrów.



(a)



(b)

Symulacja GA

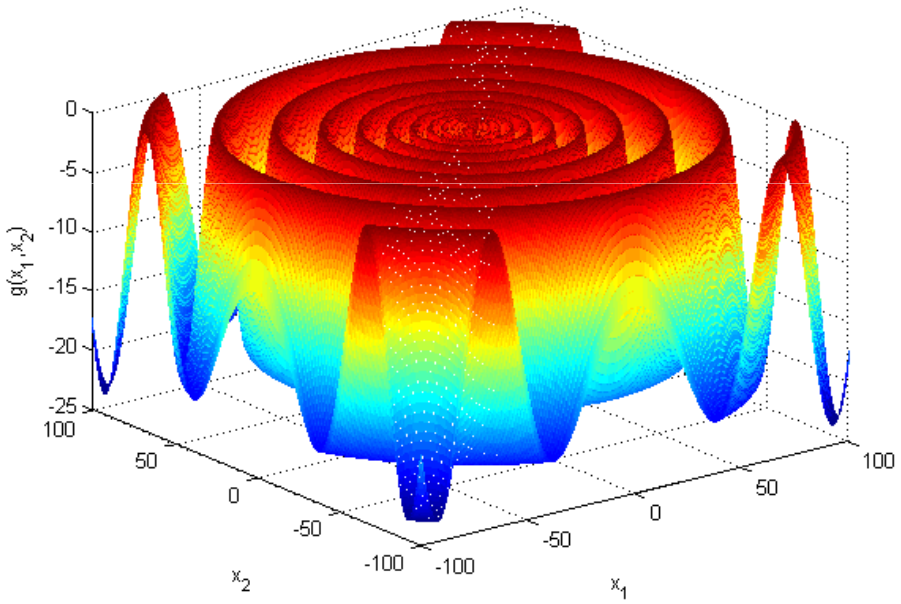
- Maksymalizacja odwróconej funkcji Schaffer'a

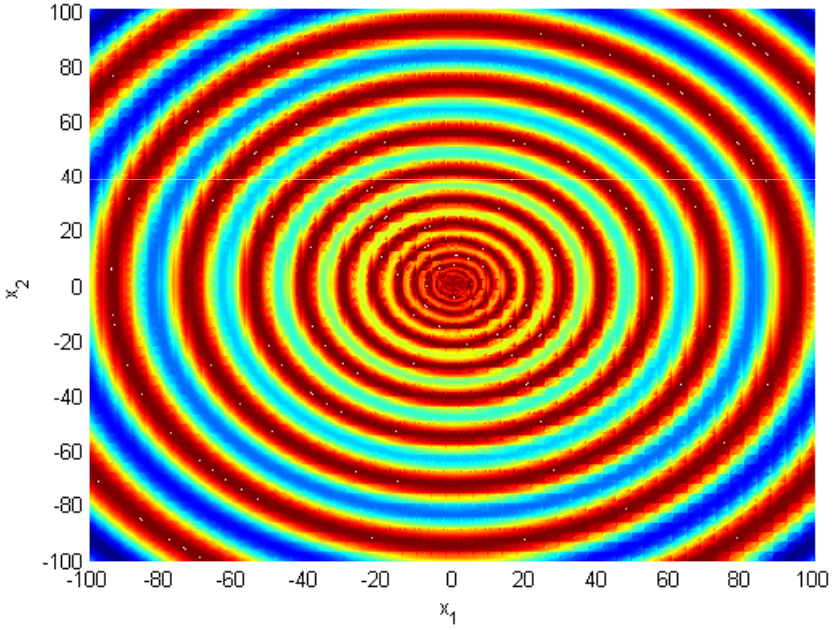
$$\max_{x_1,x_2} g(x_1,x_2) = (x_1^2 + x_2^2)^{0.25}[\sin(50(x_1^2 + x_2^2)^{0.1}) + 1]$$

- Globalne maksimum w punkcie $\begin{bmatrix} x_1^* & x_2^* \end{bmatrix} = [0 \quad 0]$ o wartości funkcji $g(0,0) = 0$

Symulacja GA

Odwrócona funkcja Schaffer’a

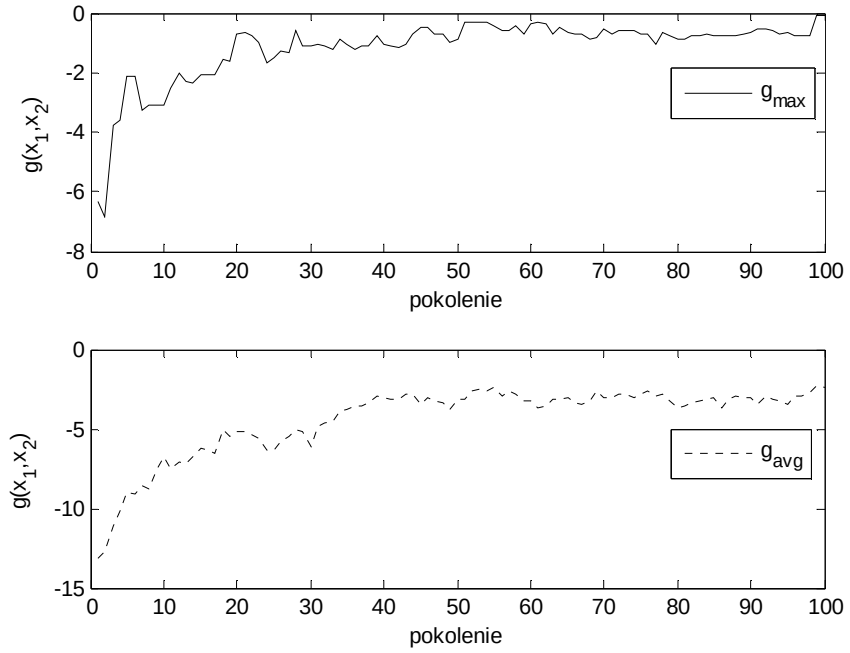




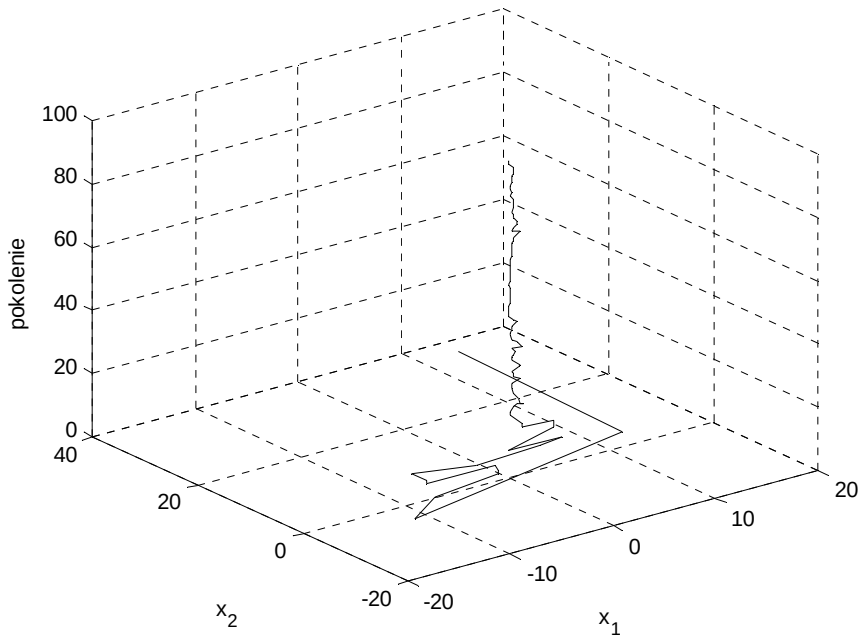
Symulacja GA

Przebiegi algorytmu genetycznego dla:

a) maksymalnej i średniej wartości funkcji zysku
b) poszukiwanych parametrów.



(a)



(b)

Symulacja GA

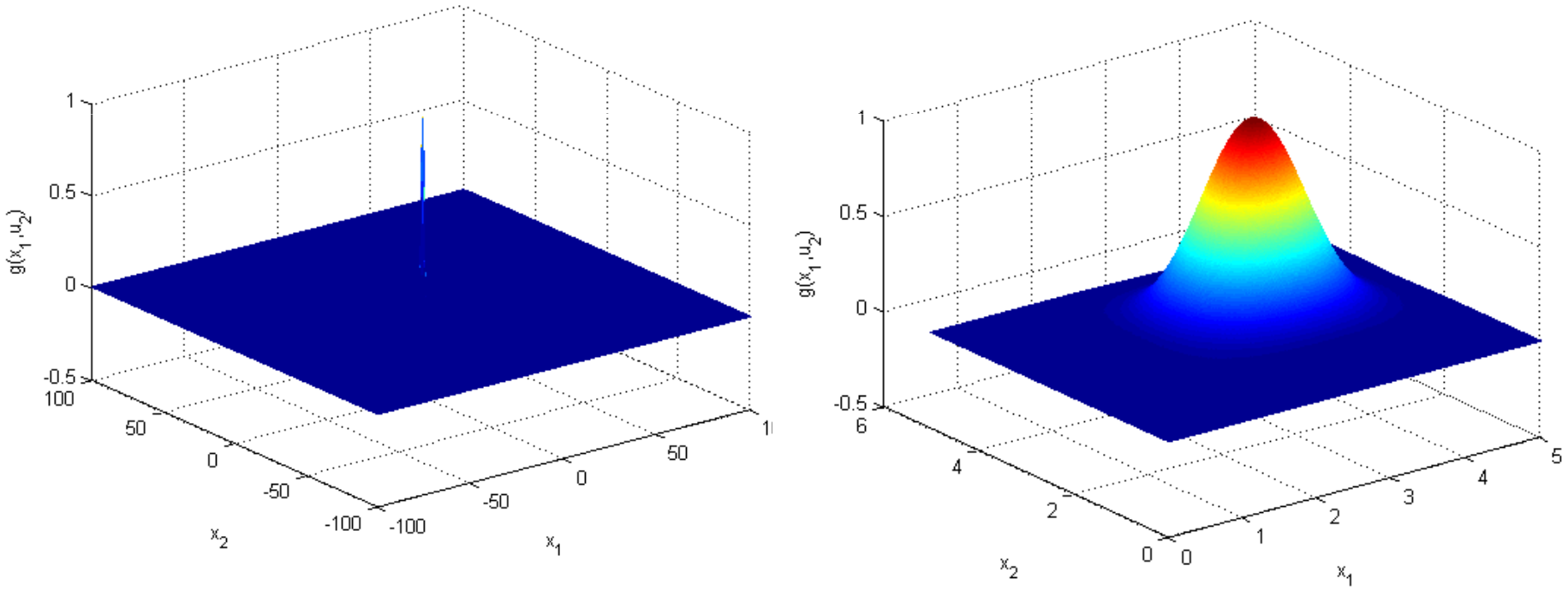
- Maksymalizacja odwróconej funkcji Easom'a

$$\max_{x_1, x_2} g(x_1, x_2) = \cos(x_1)\cos(x_2)\exp\left(-(x_1-\pi)^2-(x_2-\pi)^2\right)$$

- Globalne maksimum w punkcie $\begin{bmatrix} x_1^* & x_2^* \end{bmatrix} = \begin{bmatrix} \pi & \pi \end{bmatrix}$ o wartości funkcji $g(\pi,\pi) = 1$

Symulacja GA

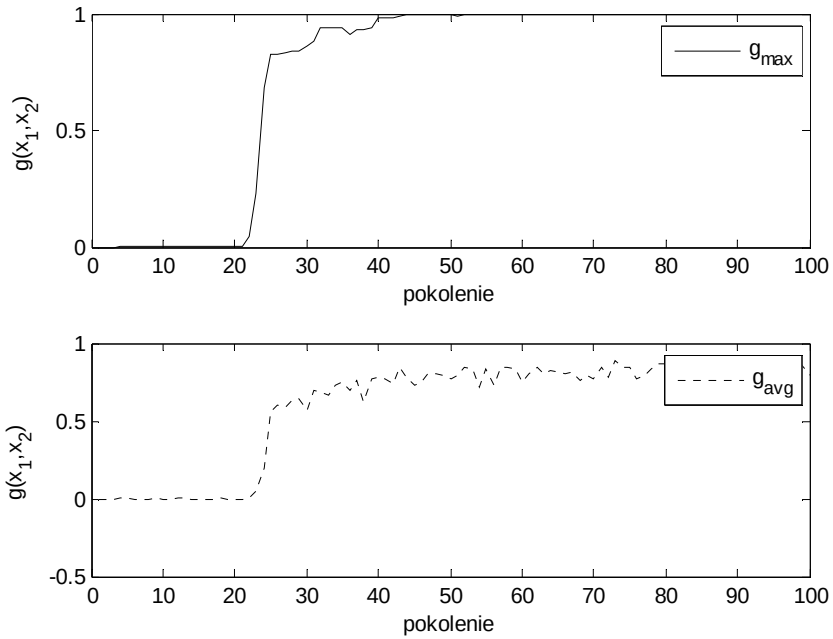
Odwrócona funkcja Easom'a



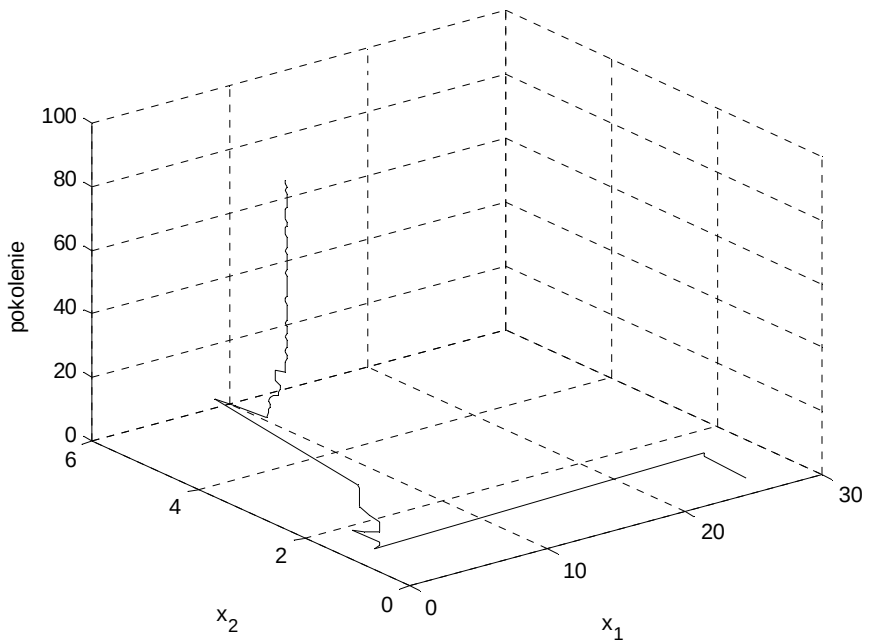
Symulacja GA

Przebiegi algorytmu genetycznego dla:

a) maksymalnej i średniej wartości funkcji zysku
b) poszukiwanych parametrów.



(a)



(b)

Symulacja GA

- Maksymalizacja funkcji wielomodalnej

$$\max_{x_1,\, x_2} g(x_1, x_2) = \sum_{j=1}^{5} \left((x_1 - a_{1j})^2 + (x_2 - a_{2j})^2 + c_j \right)^{-1}$$

gdzie

$$a_{11} = 4,\ a_{12} = 1,\ a_{13} = 8,\ a_{14} = 6,\ a_{15} = 7$$

$$a_{21} = 4,\ a_{22} = 1,\ a_{23} = 8,\ a_{24} = 6,\ a_{25} = 3$$

$$c_1 = 0.1,\ c_2 = 0.2,\ c_3 = 0.2,\ c_4 = 0.4,\ c_5 = 0.6$$

Symulacja GA

- Rozważana funkcja kryterialna posiada jedno globalne oraz cztery lokalne maksima:

$$g(4,4) \approx 10.2994$$

$$g(8,8) \approx 5.1980$$

$$g(1,1) \approx 5.1099$$

$$g(6,6) \approx 2.8597$$

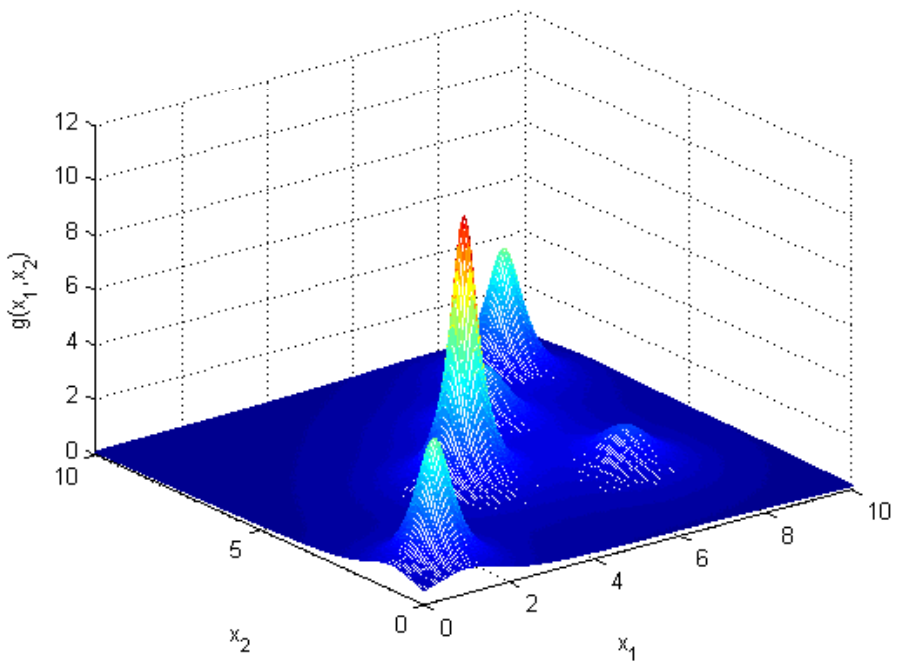
$$g(7,3) \approx 1.9249$$

Symulacja GA

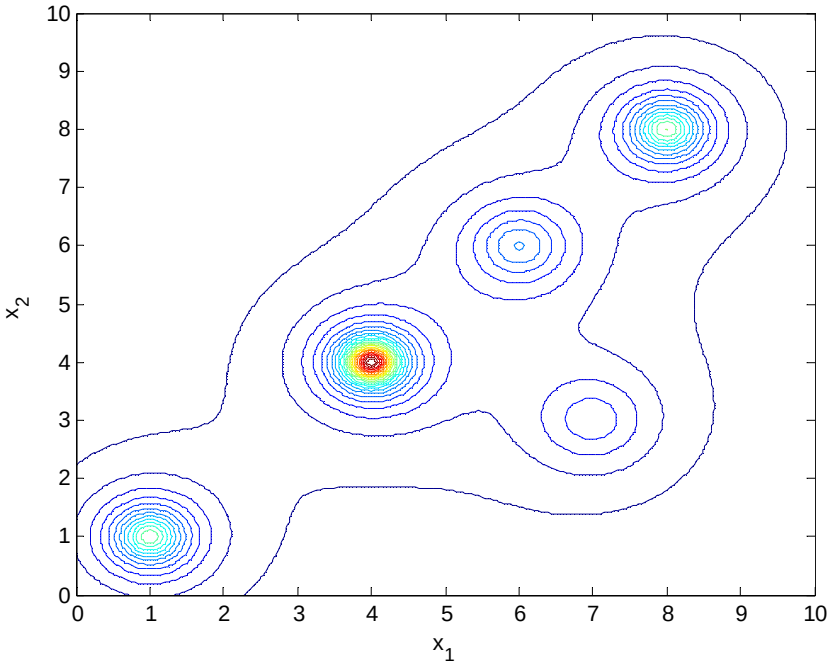
Odwrócona funkcja Shekel'a:

a) wykres

b) zrzutowana na płaszczyznę poszukiwanych parametrów



(a)



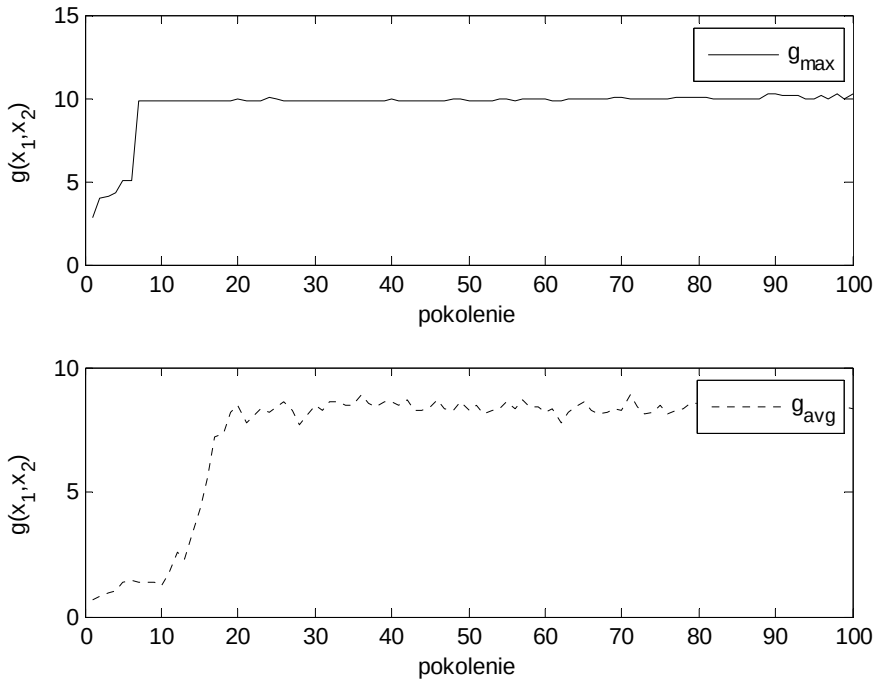
(b)

Symulacja dla klasycznego GA

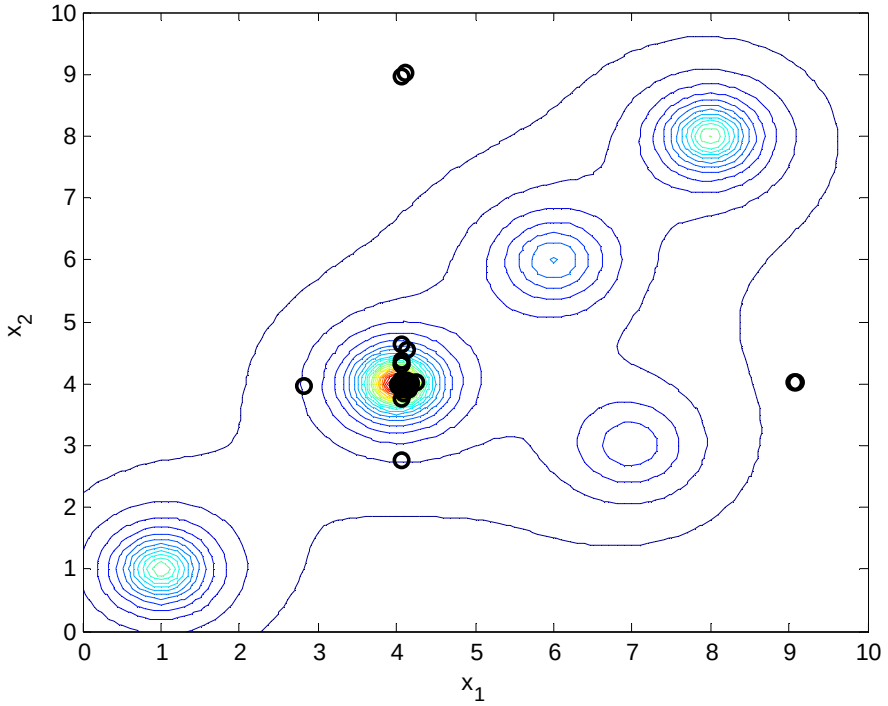
Przebiegi algorytmu genetycznego dla:

a) maksymalnej i średniej wartości funkcji zysku

b) populacja osobników w 100 pokoleniu na tle płaszczyzny poszukiwanych parametrów



(a)



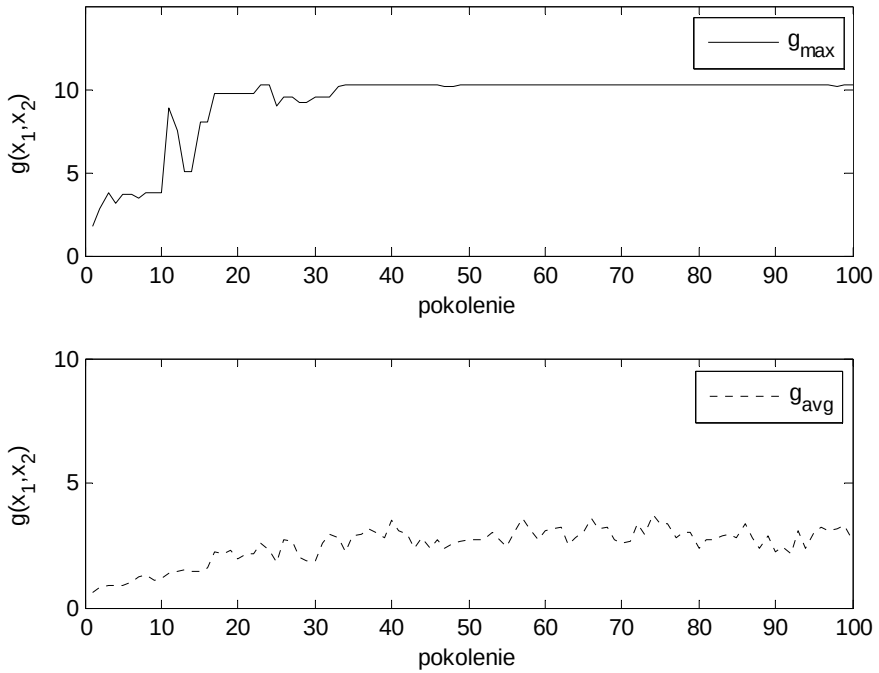
(b)

Symulacja dla GA z niszowaniem

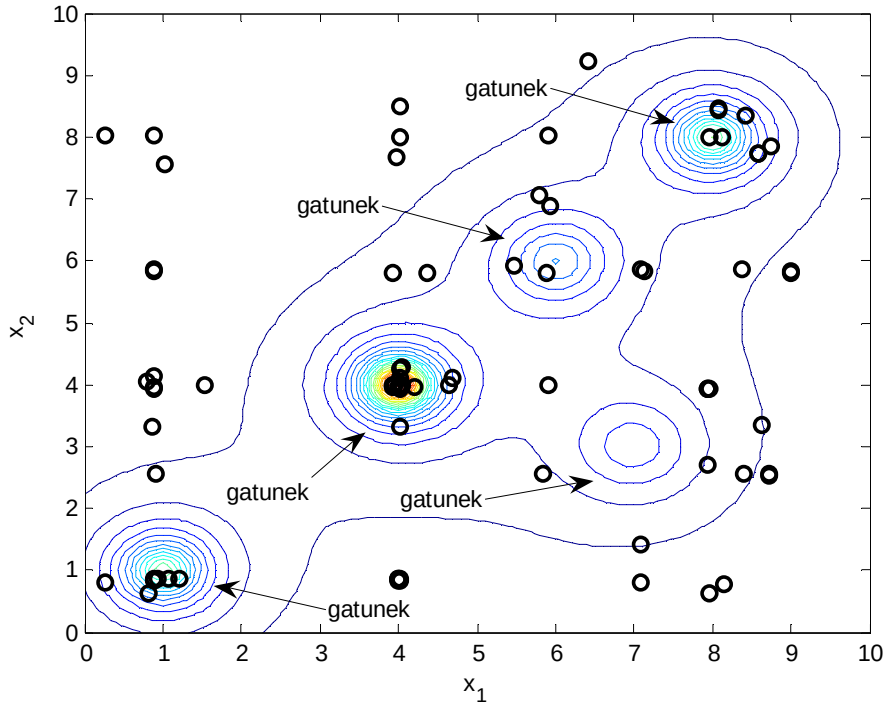
Przebiegi algorytmu genetycznego z zastosowanym niszowaniem dla:

a) maksymalnej i średniej wartości funkcji zysku

b) populacja osobników w 100 pokoleniu na tle płaszczyzny poszukiwanych parametrów



(a)



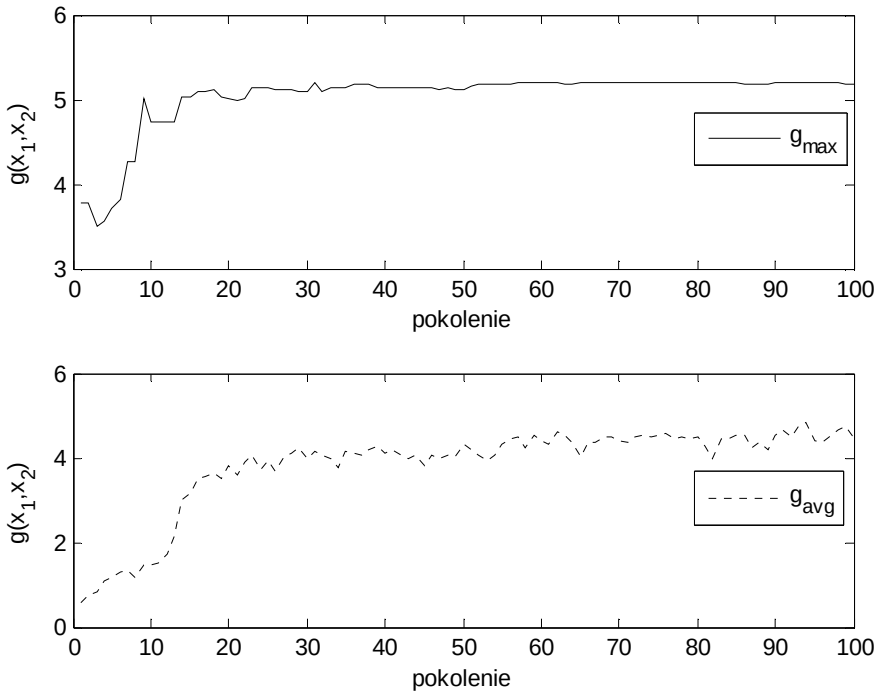
(b)

Symulacja dla klasycznego GA (mały rozmiar populacji)

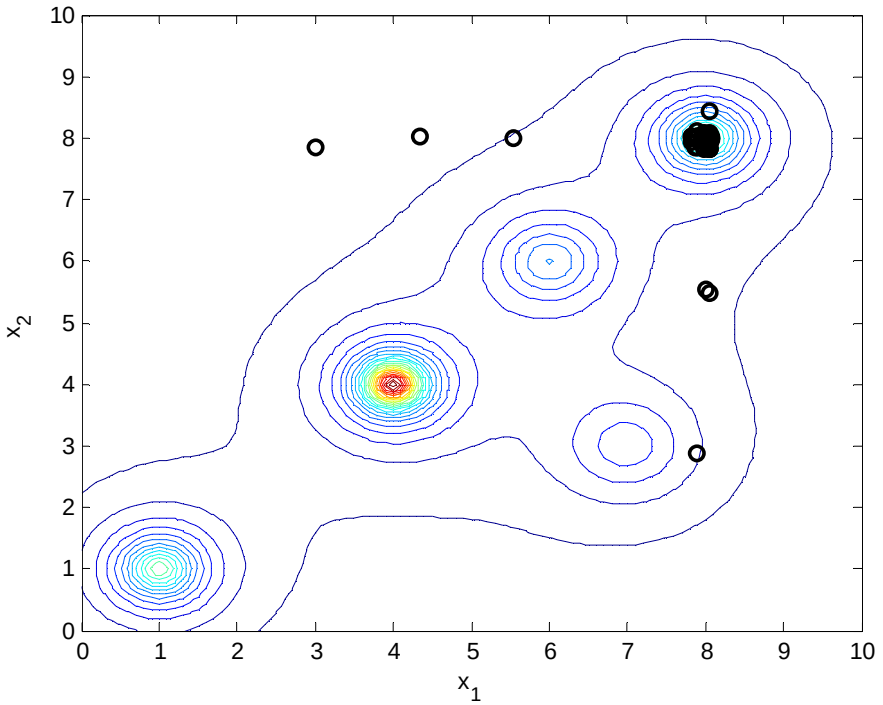
Przebiegi algorytmu genetycznego dla:

a) maksymalnej i średniej wartości funkcji zysku

b) populacja osobników w 100 pokoleniu na tle płaszczyzny poszukiwanych parametrów



(a)



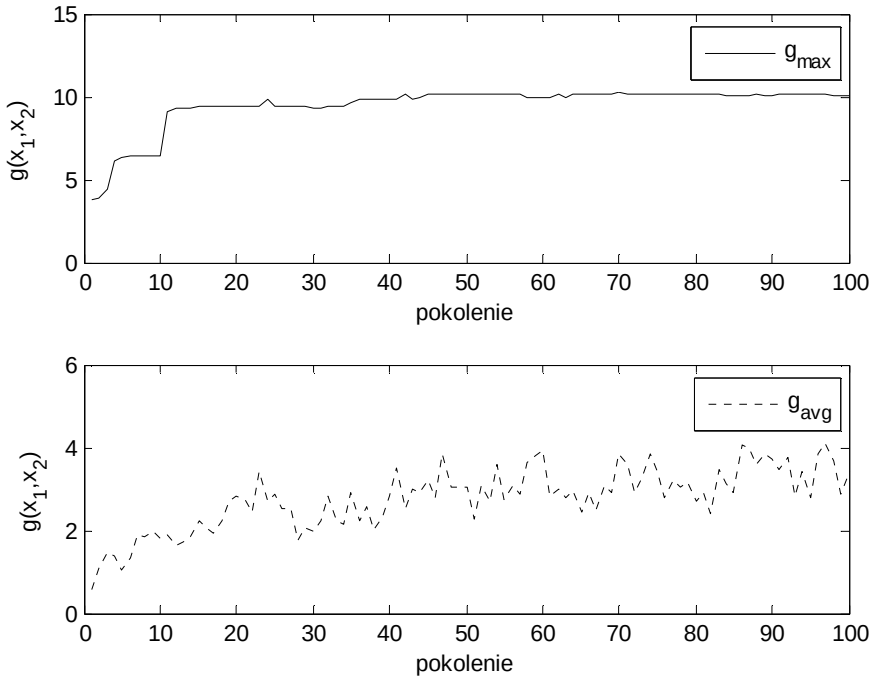
(b)

Symulacja dla GA z niszowaniem (mały rozmiar populacji)

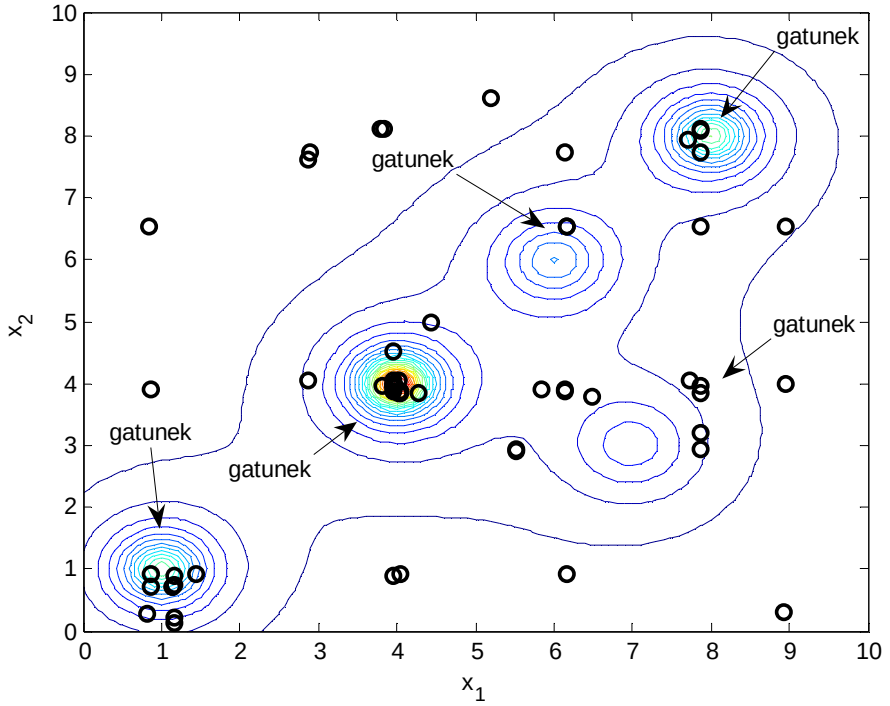
Przebiegi algorytmu genetycznego z zastosowanym niszowaniem dla:

a) maksymalnej i średniej wartości funkcji zysku

b) populacja osobników w 100 pokoleniu na tle płaszczyzny poszukiwanych parametrów



(a)



(b)

Strategie Ewolucyjne (SE)

- SE powstały w wyniku badań optymalizacyjnych urządzeń mechanicznych metodą permutowania pewnego początkowego rozwiązania
- Rodzaje strategii ewolucyjnych:
  - (1+1)
  - $(\mu+\lambda)$
  - (μ, λ)

Strategia ewolucyjna (1+1)

- Przetwarzany jeden osobnik haploidalny z kodowaniem multiallelicznym
- Osobnik jest modyfikowany poprzez mutację fentypową z rozkładem normalnym
- Rodzic zastępowany potomkiem, jeżeli przystosowanie potomka jest lepsze
- Mutacja wykorzystuje mechanizm adaptacji jej zasięgu zwany **regułą 1/5 sukcesów**

Strategia ewolucyjna (1+1)

- Reguła 1/5 sukcesów:
  - zwiększa zasięg mutacji, jeżeli pewna założona liczba sukcesów mutacji jest większa niż 1/5 ogólnej liczby wykonywanych mutacji
  - w przeciwnym przypadku zasięg mutacji jest odpowiednio zmniejszony
- SE (1+1) podobna do algorytmu symulowanego wyżarzania, ale nie dopuszcza się akceptacji punktu gorszego

Schemat SE (1+1)

1. wygeneruj losowo z rozkładem równomiernym początkowego osobnika

$$\boldsymbol{v} = \boldsymbol{x} = [x_1 \quad x_2 \quad \ldots \quad x_n]^{\mathrm{T}} \in \Re^n$$

2. wyznacz stopień przystosowania $f(\boldsymbol{x}) \in \Re$
3. jeżeli liczba iteracji SE (1+1) osiągnęła wartość założoną to zakończ poszukiwania i wyznacz najlepiej przystosowanego osobnika; w przeciwnym przypadku przejdź do kroku 4

Schemat SE (1+1)

4. utwórz potomka poprzez operację mutacji każdego genu osobnika

$$x_j' = x_j + r \cdot \varepsilon$$

gdzie

x_j' – zmodyfikowany *j*-ty gen osobnika

ε – liczba losowa wg rozkładu normalnego

r – zasięg mutacji (początkowa wartość r=1)

Schemat SE (1+1)

5. wyznacz stopień przystosowania potomka

$$f(\boldsymbol{x}') \in \Re$$

6. jeżeli $f(\boldsymbol{x}') > f(\boldsymbol{x})$ to zastąp rodzica potomkiem i zwiększ o jeden licznik sukcesów mutacji *s* (początkowa wartość *s*=0)

$$s = s + 1$$

Schemat SE (1+1)

7. zastosuj regułę 1/5 sukcesu
   - jeżeli $s > 1/5$ wszystkich sukcesów to zwiększ zasięg mutacji wg wzoru $r = (0.82)^{-1} \cdot \varepsilon$
   - jeżeli $s = 1/5$ wszystkich sukcesów to zasięg mutacji pozostaw bez zmian
   - jeżeli $s < 1/5$ wszystkich sukcesów to zmniejsz zasięg mutacji wg wzoru $r = 0.82 \cdot \varepsilon$
8. powróć do kroku 3

Podsumowanie SE (1+1)

- Wartość 1/5 oraz pozostałe współczynniki modyfikacji zasięgu mutacji zostały wyznaczone eksperymentalnie

- Mimo zastosowanego mechanizmu adaptacji mutacji strategię ewolucyjną cechuje niewielka odporność na minima lokalne

Strategia Ewolucyjna (μ+λ)

- Uogólnienie SE (1+1) przetwarzana populacje μ osobników o strukturze diploidalnej

$$\boldsymbol{v}_{di} = \left[\boldsymbol{v}_{1_i} \quad \boldsymbol{v}_{2_i}\right]^{\mathrm{T}}, \quad i = 1, 2, \ldots, N$$

gdzie

$$\boldsymbol{v}_{1_i} = \left[x_{1_i} \quad x_{2_i} \quad \ldots \quad x_{n_i}\right]^{\mathrm{T}} \in \Re^n$$

$$\boldsymbol{v}_{2_i} = \left[\sigma_{1_i} \quad \sigma_{2_i} \quad \ldots \quad \sigma_{n_i}\right]^{\mathrm{T}} \in \Re^n$$

Strategia Ewolucyjna (μ+λ)

- Chromosom $v_{1_i} = x_i$ jest wektorem poszukiwanych parametrów
- Chromosom v_{2_i} jest wektorem odchyleń standardowych dla każdego parametru (wykorzystywany przy operacji mutacji)
- Selekcja osobników polega na wylosowaniu (ze zwracaniem) z rozkładem równomiernym λ osobników spośród μ osobników

Strategia Ewolucyjna (μ+λ)

- Rodzice poddawani są procesowi krzyżowania arytmetycznego oraz mutacji
- Potomkowie porównywani są ze z rodzicami
- Do dalszego cyklu strategii wybieranych jest tylko μ najlepszych osobników (elityzm)

Schemat SE (μ+ λ)

1. wygeneruj losowo z rozkładem równomiernym populację μ osobników

$$\boldsymbol{V}_d^{\mu} = \left[\boldsymbol{v}_{d1} \quad \boldsymbol{v}_{d2} \quad \cdots \quad \boldsymbol{v}_{d\mu}\right]$$

2. wyznacz przystosowanie $f(\boldsymbol{x}_i) \in \Re$ osobników

3. jeżeli liczba iteracji SE (μ+λ) osiągnęła wartość założoną to zakończ poszukiwania i wyznacz najlepiej przystosowanego osobnika; w przeciwnym przypadku przejdź do kroku 4

Schemat SE (μ+ λ)

4. utwórz pulę rodzicielską poprzez λ-losowań ze zwracaniem osobnika z populacji μ-osobników
5. utwórz z puli rodzicielskiej populację $\boldsymbol{V}_d^{\lambda}$ potomków poprzez:
   - przeprowadzenie krzyżowania arytmetycznego dla wylosowanych par rodziców z populacji $\boldsymbol{V}_d^{\lambda}$

Schemat SE (μ+ λ)

– przeprowadzenie dla $\boldsymbol{v}_{di}$ mutacji

$$\boldsymbol{v}'_{2_i} = \left[\sigma_{1_i} e^{(\tau\zeta+\upsilon\zeta_1)} \quad \ldots \quad \sigma_{n_i} e^{(\tau\zeta+\upsilon\zeta_n)}\right]^{\mathrm{T}} \in \Re^n$$

$$\tau = \left(\sqrt{2\sqrt{n}}\right)^{-1} \qquad \upsilon = \left(\sqrt{2n}\right)^{-1}$$

$$\boldsymbol{v}'_{1_i} = \left[x_{1_i} + \sigma'_{1_i}\varepsilon_1 \quad \ldots \quad x_{n_i} + \sigma'_{n_i}\varepsilon_n\right]^{\mathrm{T}} \in \Re^n$$

gdzie

ε_i, ζ_i, ζ - wartości wygenerowane losowo z rozkładem normalnym

Schemat SE (μ+ λ)

6. wyznacz stopień przystosowania potomków

$$f(\boldsymbol{x}_i') \in \Re$$

7. utwórz nową populacje μ najlepszych osobników spośród populacji $\boldsymbol{V}_d^{\lambda}$ $\boldsymbol{V}_d^{\mu}$

8. powróć do kroku 3

Podsumowanie SE (μ+ λ)

- Mutacja przeprowadzana jest w pierwszej kolejności dla chromosomu z odchyleniami standardowymi
- Zmodyfikowany wektor z odchyleniami standardowymi jest wykorzystywany przy mutacji chromosomu reprezentującego poszukiwane parametry zadania optymalizacji
- Mechanizm mutacji ma na celu zapewnienie samoczynnej adaptacji zasięgu mutacji

Podsumowanie SE (μ+ λ)

- Brak arbitralności przy doborze poszczególnych odchyleń standardowych (dobierane losowo)
- Efekt adaptacji jest wynikiem zachodzącego procesu selekcji wybierającego do dalszej ewolucji lepiej przystosowane osobniki
- SE (μ+ λ) pomimo większej odporności od (1+1) również narażona jest na zjawisko przedwczesnej zbieżności algorytmu

Strategia Ewolucyjna (μ, λ)

- Udoskonalona wersja strategii (μ+λ)
- Mniej narażona na zjawisko przedwczesnej zbieżności poprzez zrezygnowanie z mechanizmu tworzenia elity
- W odróżnieniu od (μ+λ) utworzenie nowej populacji μ najlepszych osobników odbywa się tylko dla populacji potomków $\boldsymbol{V}_d^{\lambda}$

Podsumowanie Strategii Ewolucyjnych

- W strategiach ewolucyjnych (μ+λ) i (μ, λ) parametry algorytmu zwykle przyjmują następujące wartości:
  - $\boldsymbol{V}_d^{\mu}$ złożona z $\mu \in [50, 100]$ osobników
  - $\boldsymbol{V}_d^{\lambda}$ zawierająca $\lambda = 4 \div 6$ razy więcej osobników niż populacja $\boldsymbol{V}_d^{\mu}$

Programowanie ewolucyjne (PE)

- Problem stworzenia gramatyki nieznanego języka na podstawie zestawu symboli i wyrażeń syntaktycznie poprawnych
- Gramatyka modelowana za pomocą poszukiwanych w sposób ewolucyjny grafów złożonych ze stanów i funkcji przejść
- Proces tworzenia nowego grafu poprzez mutację jego struktury (dodanie/usunięcie stanu, zmianę funkcji wyjść lub przejść)

Programowanie ewolucyjne

- W późniejszych badaniach programowanie ewolucyjne przeszło w stronę strategii ewolucyjnych jako kolejnych metod optymalizacji numerycznej
- Podobnie jak w SE $(\mu+\lambda)$ i (μ, λ) osobnik w PE jest reprezentowany kodowaniem multiallelicznym (zminennopozycyjnym)
- W odróżnieniu od SE, każdy osobnik generuje dokładnie jednego potomka

Programowanie ewolucyjne

- Tworzenie potomków przeprowadzane wg mutacji fenotypowej w trzech wariantach:
  - **bez adaptacji**
  - **adaptacja ze śledzeniem wartości funkcji przystosowania**
  - **samoczynna adaptacja**

- Struktura osobnika haploidalna lub diploidalna w zależności od stosowanej mutacji

PE – mutacja bez adaptacji

- Modyfikacja osobnika haploidalnego (z pojedynczym chromosomem) wg wzoru

$$\boldsymbol{v}_i^{'} = \boldsymbol{x}_i^{'} = \left[x_{1_i} + \sigma_1 \varepsilon_1 \quad \ldots \quad x_{n_i} + \sigma_n \varepsilon_n \right]^{\mathrm{T}} \in \Re^n$$

gdzie

ε - wektor zmiennych losowych o *n*-wymiarowym nieskolerowanym rozkładzie normalnym

σ - zasięg mutacji

PE – mutacja adaptacyjna śledząca zmiany przystosowania

- Generowanie potomka haploidalnego

$$\boldsymbol{v}_i' = \boldsymbol{x}_i' = \boldsymbol{x}_i + \varepsilon\sqrt{\beta(M - f(\boldsymbol{x}_i)) + \gamma} \in \Re^n$$

gdzie

ε - wektor zmiennych losowych o *n*-wymiarowym nieskolerowanym rozkładzie normalnym

γ, β - parametry PE (zwykle β=1, γ=0)

M - oszacowanie wartości globalnego maksmimum

PE – mutacja z samoczynną adaptacją

- Stosowana dla osobników diploidalnych
- W pierwszej kolejności modyfikowany jest chromosom zawierający poszukiwane parametry zadania optymalizacji
- W drugi kroku zmieniany jest chromosom składający się z genów będącymi wartościami odchyleń standardowych poszczególnych parametrów chromosomu

PE – mutacja z samoczynną adaptacją

- Potomek uzyskiwany jest wg wzorów

$$\boldsymbol{v}'_{1_i} = \left[x_{1_i} + \sigma_1 \varepsilon_1 \quad \ldots \quad x_{n_i} + \sigma_n \varepsilon_n \right]^{\mathrm{T}}$$

$$v'_{2\,j_i} = \sqrt{\min\{\sigma^2_{min},\ (\sigma_1^2 + \zeta \varepsilon_1)\}}$$

gdzie

$v'_{2\,j_i}$ - j-ty gen chromosomu $\boldsymbol{v}'_{2_i}$

ε_j - zmienna losowa o rozkładzie normalnym

σ_j - j-te odchylenie standardowe

ζ - intensywność zmian σ_j

Programowanie ewolucyjne

- Strategia podstawienia wg rang osobników
- Do nowej populacji wybierani są osobnicy o najwyższej wartość rangi
- Ranga osobnika wyznaczana poprzez ‘turniej’ z wybraną losowo z grupą *k* osobników
- Wartość rangi osobnika równa liczbie osobników o przystosowaniu mniejszym od przystosowania danego osobnika

Schemat PE

1. wygeneruj losowo z rozkładem równomiernym populację μ osobników

$$\boldsymbol{V}_d^{\mu} = \begin{bmatrix} \boldsymbol{v}_{d1} & \boldsymbol{v}_{d2} & \cdots & \boldsymbol{v}_{d\mu} \end{bmatrix}$$

2. wyznacz przystosowanie $f(\boldsymbol{x}_i) \in \Re$ osobników

3. jeżeli liczba iteracji PE osiągnęła wartość założoną to zakończ poszukiwania i wyznacz najlepiej przystosowanego osobnika; w przeciwnym przypadku przejdź do kroku 4

Schemat PE

4. utwórz populację $\boldsymbol{v}_d^{\lambda}$ z populacji $\boldsymbol{v}_d^{\mu}$ ($\mu = \lambda$) poprzez mutację z samoczynną adaptacją

$$\boldsymbol{v}'_{1_i} = \left[x_{1_i} + \sigma_1 \varepsilon_1 \quad \ldots \quad x_{n_i} + \sigma_n \varepsilon_n \right]^{\mathrm{T}}$$

$$v'_{2\,j_i} = \sqrt{\min\{\sigma_{min}^2,\ (\sigma_1^2 + \zeta \varepsilon_1)\}}$$

5. wyznacz stopień przystosowania każdego potomka

Schemat PE

6. nadaj rangę osobnikom z populacji $\boldsymbol{V}_d^{\lambda}$ i $\boldsymbol{V}_d^{\mu}$:
   - wylosowanie ze zwracaniem dla osobnika $\boldsymbol{v}_{di}$ grupy osobników z populacji $\boldsymbol{V}_d^{\lambda}$ i $\boldsymbol{V}_d^{\mu}$
   - ranga osobnika $\boldsymbol{v}_{di}$ równa liczbie osobników z wybranej losowo grupy, dla których przystosowanie jest mniejszy od $f(\boldsymbol{v}_{di})$
7. zastąp populację $\boldsymbol{V}_d^{\mu}$ osobnikami z najwyższą rangą
8. powróć do kroku 3.