

# **Algorytm Genetyczny**

zastosowanie do procesów rozmieszczenia stacji raportujących w sieciach komórkowych

# Dlaczego Algorytmy Inspirowane Naturą?

---

Rozwój nowych technologii: złożone problemy obliczeniowe w nauce i technice XX i XXI wieku:

- ▶ duże przestrzenie potencjalnych rozwiązań
- ▶ rozwiązania suboptymalne są zadowalające dla wielu zastosowań
- ▶ jak odnaleźć w rozsądnym czasie dobre (suboptymalne) rozwiązanie

Tradycyjne metody poszukiwania rozwiązań realnych problemów były mało efektywne

Konieczność opracowania nowych algorytmów przeszukiwania przestrzeni rozwiązań:

- ▶ mechanizmy Natury i biologii są bardzo obiecujące



# Złożoność Problemów

---

**Problem klasy P** (ang. *deterministic polynomial* - deterministycznie wielomianowy) to problem decyzyjny, dla którego rozwiązanie można znaleźć w czasie wielomianowym.

**Problem klasy NP** (niedeterministycznie wielomianowy, ang. *nondeterministic polynomial*) to problem decyzyjny, dla którego rozwiązanie można zweryfikować w czasie wielomianowym.

Różnica pomiędzy problemami **P** i **NP** polega na tym, że w przypadku **P** **znalezienie** rozwiązania ma mieć złożoność wielomianową, podczas gdy dla **NP** **sprawdzenie** podanego z zewnątrz rozwiązania ma mieć taką złożoność.

W teorii złożoności obliczeniowej **problem NP-trudny (NPH)** to taki problem obliczeniowy, którego rozwiązanie jest co najmniej tak trudne jak rozwiązanie każdego problemu z klasy **NP**.

---



# Literatura nt.

## algorytmów inspirowanych Naturą

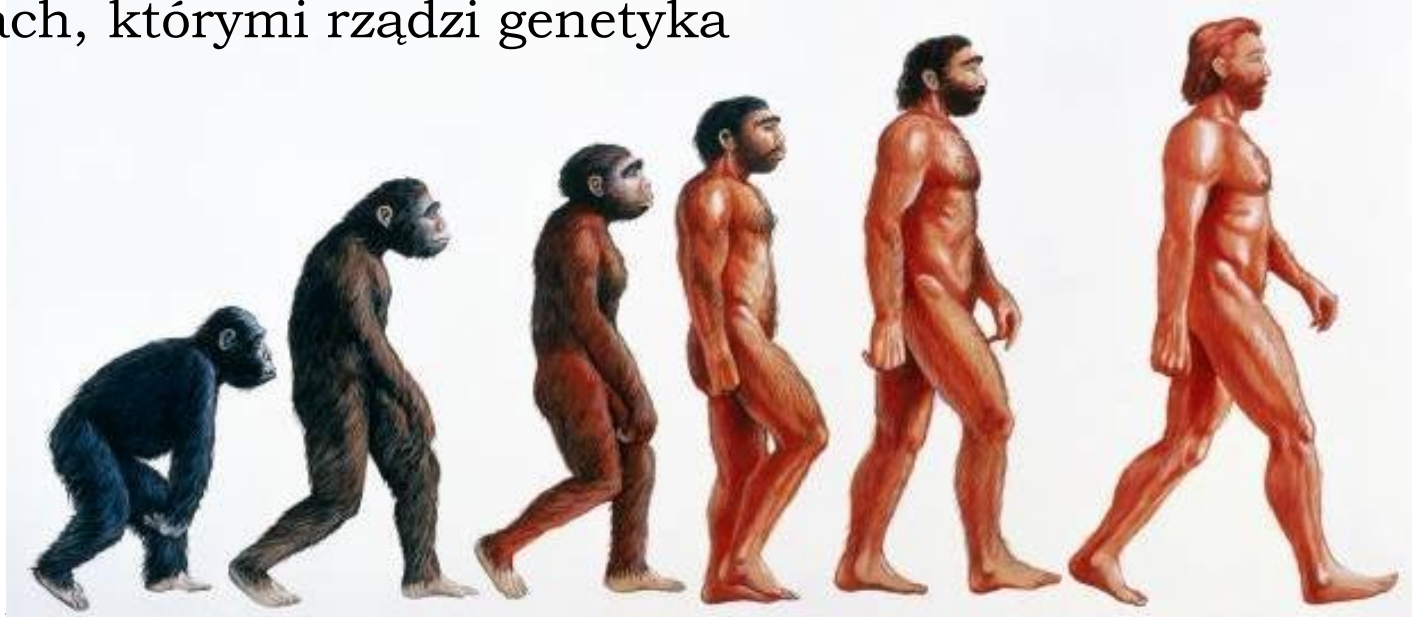
---

1. Z. Michalewicz, *Algorytmy genetyczne + struktury danych = programy ewolucyjne*, WNT, 2003
  2. L. Rutkowski, *Metody i techniki sztucznej inteligencji*, PWN, 2009
  3. Z. Michalewicz, D. B. Fogel, *Jak to rozwiązać czyli nowoczesna heurystyka*, WNT, 2006
  4. David A. Goldberg, *Algorytmy genetyczne i ich zastosowania*, WNT, 2003
  5. J. Arabas, *Wykłady z algorytmów ewolucyjnych*, WNT, 2001
  6. F. Seredyński, *Evolutionary Paradigms*
  7. Materiały z dorocznych międzynarodowych konferencji: GECCO, CEC, PPSN
- 



# Ewolucja Darwinowska

- ▶ Dzisiaj najbardziej znane algorytmy ewolucyjne są oparte na symulowanej ewolucji darwinowskiej
- ▶ **Darwinowska teoria ewolucji naturalnej** wskazuje dwa główne czynniki: **naturalną selekcję** i **genetyczne zmiany**
- ▶ **Naturalna selekcja** zakłada, że osobniki które są lepiej przystosowane do środowiska posiadają większe prawdopodobieństwo przeżycia
- ▶ Tylko te osobniki, które przeżyły mogą wydać potomstwo na zasadach, którymi rządzi genetyka




# Algorytm genetyczny

---

Twórcą algorytmów genetycznych (zaliczanych do grupy algorytmów ewolucyjnych) jest **John Henry Holland**, który z biologii czerpał inspiracje do swoich prac, dlatego sposób działania algorytmów genetycznych przypomina zjawisko ewolucji biologicznej.

Algorytmy genetyczne powstały z lat '60 na Uniwersytecie Michigan w USA, a obecnie rozwijane są przez: De Yong'a, Goldberg'a, Michalewicza i wielu innych.

---



# Terminologia algorytmów genetycznych

---

- ▶ **Populacja** to zbiór *osobników* o wspólnych cechach genetycznych.
- ▶ **Osobnikami** populacji w algorytmach genetycznych są zakodowane w postaci chromosomów zbiory *parametrów zadania*, czyli *rozwiązania*, określane też jako *punkty przestrzeni poszukiwań*. Osobniki nazywane są też *organizmami*.
- ▶ **Chromosomy** to inaczej *łańcuchy* lub *ciągi kodowe*, uporządkowane ciągi *genów*.
- ▶ **Gen** jest nazywany *cechą*, *znakiem*, *detektorem* i stanowi pojedynczy element *genotypu*, w szczególności chromosomu.
- ▶ **Genotyp**, czyli *struktura* to zespół chromosomów danego osobnika. Zatem osobnikami populacji mogą być *genotypy* albo pojedyncze chromosomy (jeśli genotyp składa się tylko z jednego chromosomu, tak się często przyjmuje).
- ▶ **Fenotyp** jest zestawem wartości odpowiadających danemu genotypowi, czyli *zdekodowaną strukturą*, a więc *zbiorem parametrów zadania* (*rozwiązaniem*, *punktem przestrzeni poszukiwań*).
- ▶ **Allel** to wartość danego *genu*, określona też jako *wartość cechy* lub *wariant cechy*.
- ▶ **Locus** to *pozycja* wskazująca miejsce położenia danego genu w *łańcuchu*, czyli chromosomie.



# Klasyczny algorytm genetyczny - Hollanda

---

- ▶ Binarne kodowanie (chromosom, łańcuch) rozwiązania problemu – osobnik populacji
- ▶ Populacja osobników – potencjalnych rozwiązań problemu
- ▶ Symulowana ewolucja z użyciem 3 operatorów stosowanych w kolejności:
  - ▶ selekcja proporcjonalna
  - ▶ krzyżowanie jednopunktowe wykonywane z prawdopodobieństwem  $p_k$
  - ▶ mutacja bitów wykonywana z prawdopodobieństwem  $p_m$





# Klasyczny algorytm genetyczny

---

Kodowanie rozwiązania problemu w postać osobnika populacji

$t = 0$

Losowa inicjalizacja populacji  $P(t)$

Ocena dopasowania osobników  $P(t)$

**WHILE** warunek stopu **NOT TRUE**

$t = t + 1$

Selekcja osobników z  $P(t - 1)$  do  $P(t)$

Krzyżowanie w  $P(t)$


Mutacja w  $P(t)$

Ocena dopasowania osobników w  $P(t)$

**DO**

Rozwiązanie problemu = najlepszy osobnik w  $P(t)$

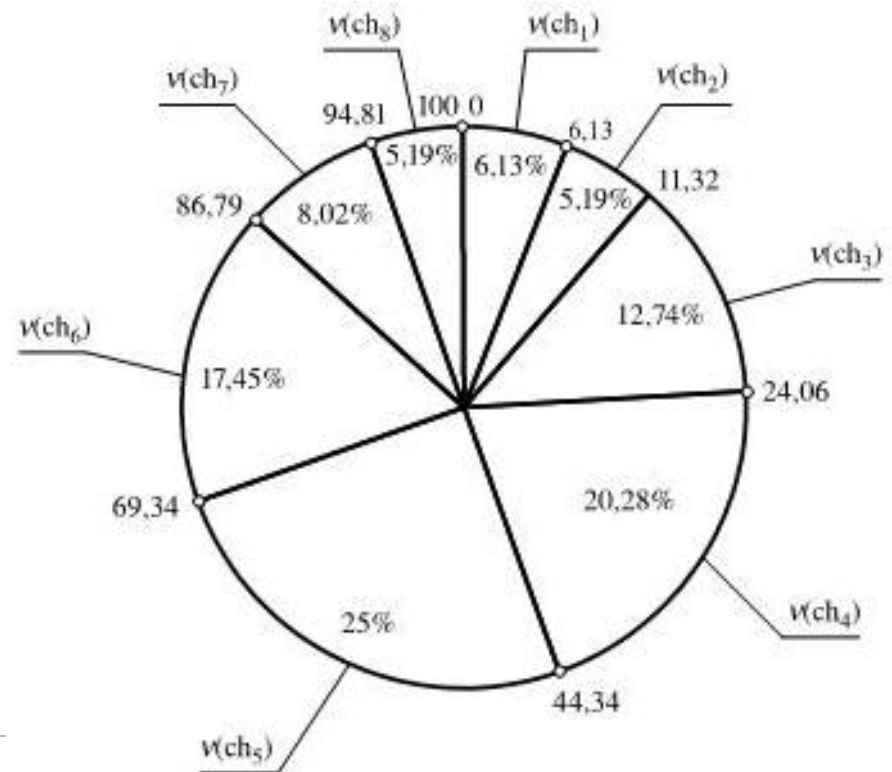
---



# Operator selekcji proporcjonalnej

- ▶ To metoda stochastycznej selekcji (symuluje naturalną selekcję), która dokonuje selekcji osobnika  $i$  (aby przeżył i miał szansę wydania potomstwa z prawdopodobieństwem:

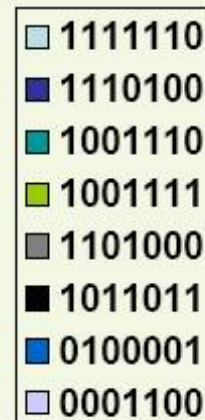
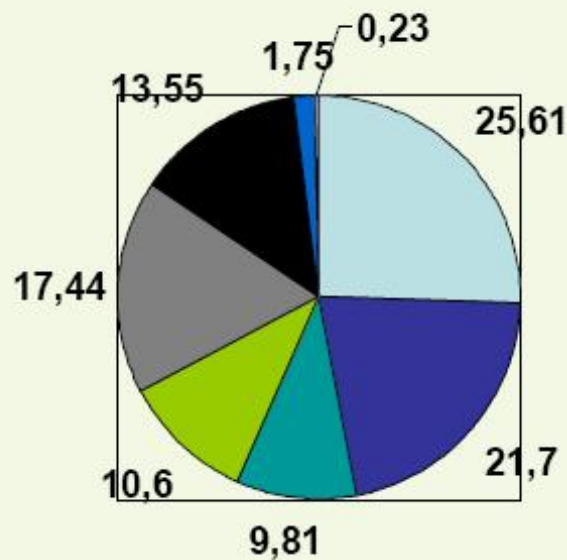
$$p_i = \frac{f_i}{\sum_{j=1}^n f_j}$$



# Selekcja ruletkowa dla funkcji dopasowania $f(x) = 2(x^2 + 1)$

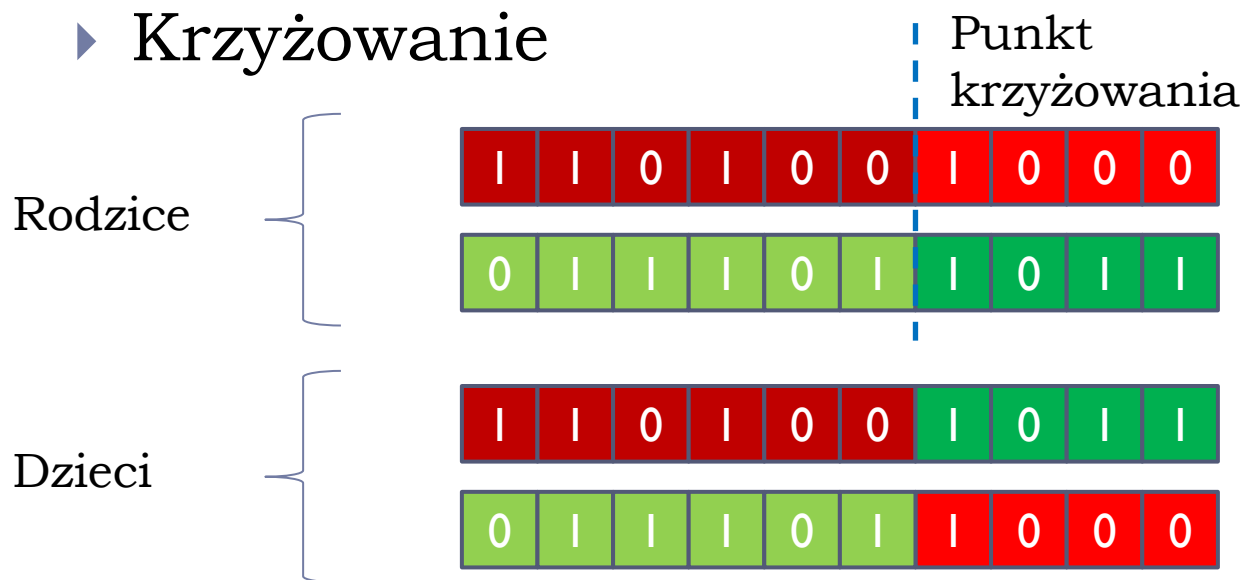
Chromosom      Dopasowanie      Procentowy udział      Rozkład na kole ruletki

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%

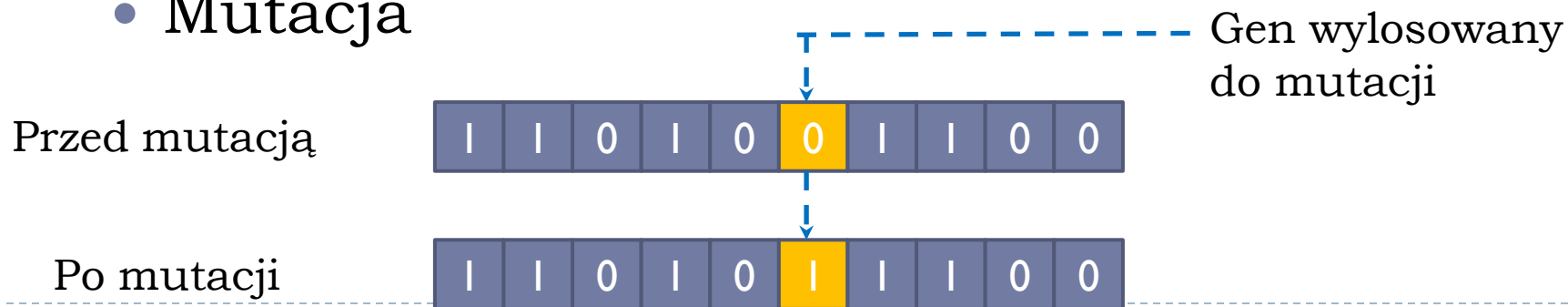


# Operatory genetyczne

## ► Krzyżowanie



## • Mutacja



# Rodzaje selekcji

---

- ▶ Proporcjonalna (koło ruletki)
- ▶ Rankingowa (rangowa)
- ▶ Turniejowa:
  - ▶ Twarda
  - ▶ Miękka
- ▶ Elitarna
- ▶ Progowa
- ▶ Stłoczenia
- ▶ ...



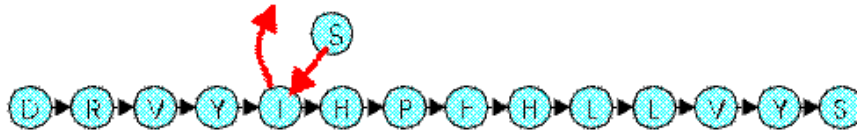
# Rodzaje krzyżowania

---

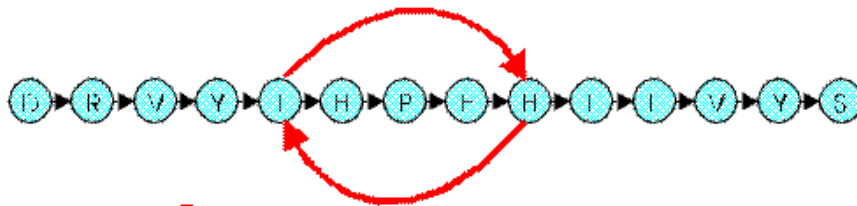
- ▶ Jednopunktowe
- ▶ Wielopunktowe
- ▶ Uśredniające
- ▶ Wg wzorca
- ▶ Jednorodne
- ▶ ...



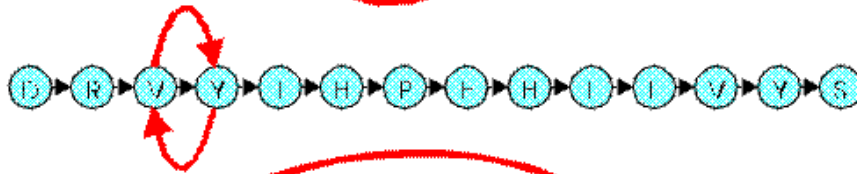
# Rodzaje mutacji



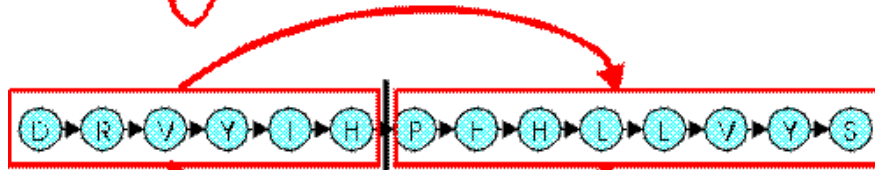
(a) Replacement



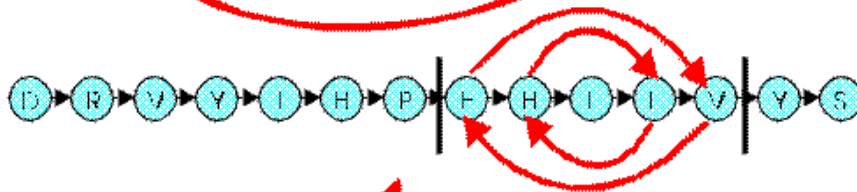
(b) Random swap



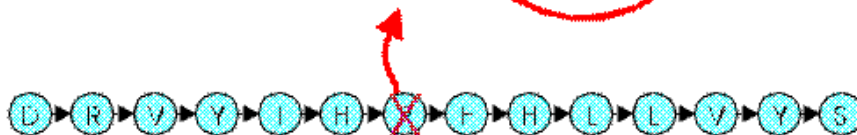
(c) Adjacent swap



(d) End-for-end swap



(e) Inversion



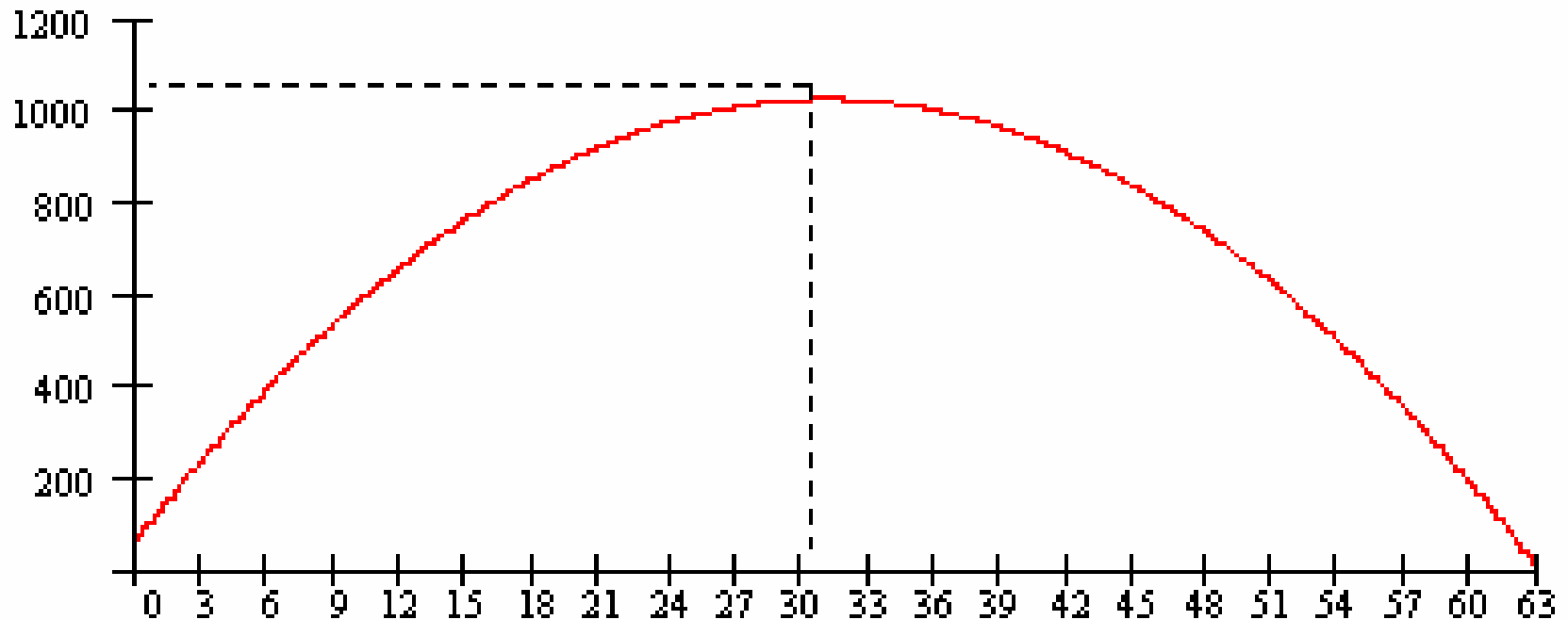
(f) Deletion



# Algorytm genetyczny – Przykład 1

---

Znaleźć maximum funkcji  $f(x) = -(x - 31)^2 + 1024$  w zakresie  $[0, 63]$  liczb całkowitych.





# Kodowanie problemu

---

- ▶ zmienna  $x$   $[0, 1, 2, \dots, 63]$  ( $x$ -liczba całkowita)
- ▶ Kodowanie binarne:  $x_{(10)} \rightarrow x_{(2)}$ ,  
np. jeżeli  $x_{(10)} = 20$ , to  $20_{(10)} \rightarrow (010100)_{(2)}$
- ▶ Aby zakodować 64 wartości zmiennej  $x$  potrzebujemy łańcucha binarnego o długości  $l = 6$
- ▶ przestrzeń rozwiązań

$$0 \rightarrow 000000$$

$$1 \rightarrow 000001$$

.

.

.

$$63 \rightarrow 111111$$



# Generowanie populacji początkowej

---

- ▶ Losowo generujemy populację składającą się z  $m$  osobników
- ▶ Jeżeli, np.  $m=6$ ,  $P(0)$  może być następująca:  
15, 35, 6, 20, 57, 0

- ▶ W binarnej formie:

001111

100011

000110

010100

111001

000000

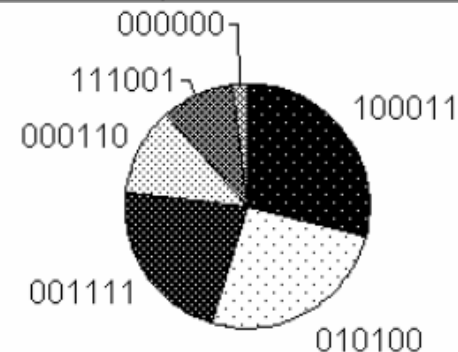


# Realizacja operatora selekcji

- ▶ Musimy obliczyć prawdopodobieństwo selekcji i-go osobnika:  
$$p_i = \frac{f_i}{\sum_{j=1}^n f_j}$$
- ▶ W tym celu należy obliczyć wartość funkcji dopasowania (fitness function) każdego osobnika w populacji

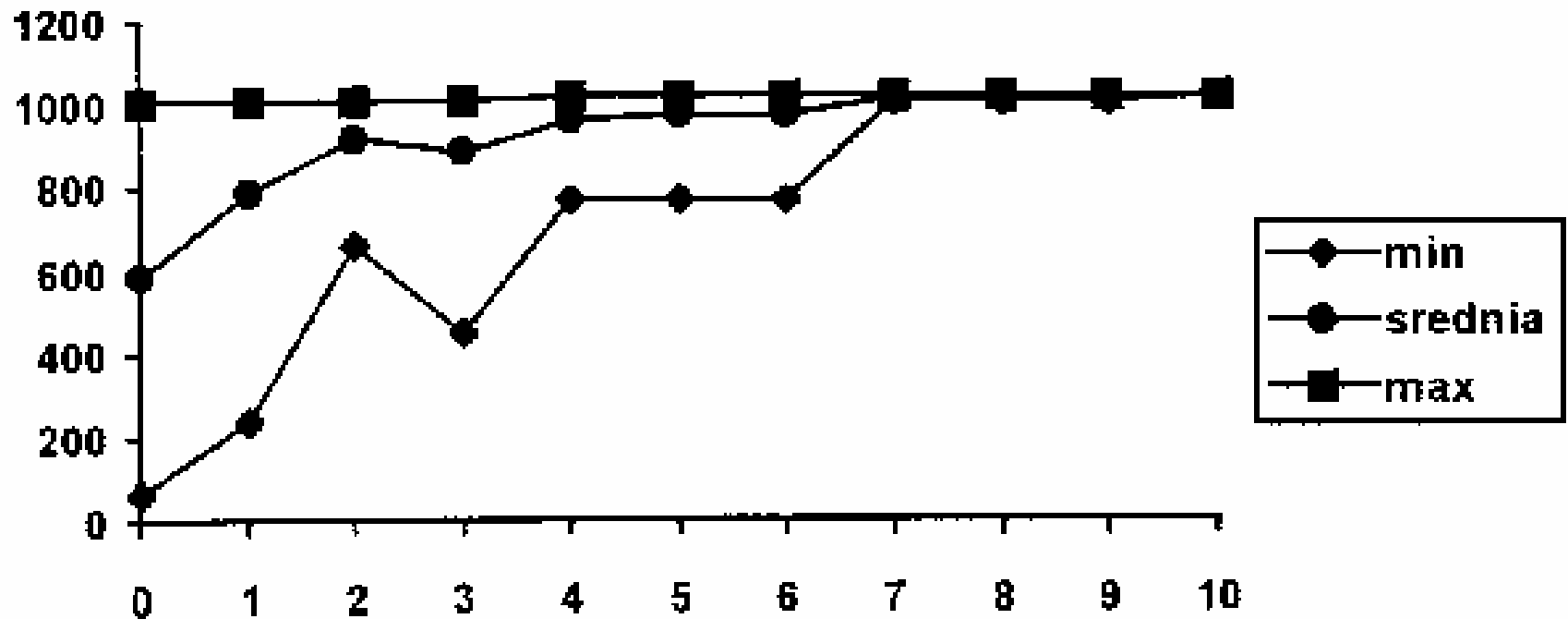
individual	Value of fitness function	% of total
100011	1008	28.89
010100	903	25.88
001111	768	22.01
000110	399	11.44
111001	348	9.97
000000	63	1.81
<u>total</u>	3489	100

- ▶ Tworzymy koło ruletki
- ▶ Operatory genetyczne (krzyżowanie i mutacja)



# Symulacja komputerowa

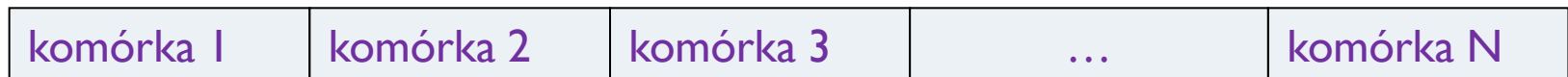
---



# Zastosowanie AG do rozmieszczenia stacji raportujących w sieciach mobilnych

Schemat kodowania:

Ilość genów w chromosomie  $L = N$ , liczbie komórek w sieci, gdzie  $N = n \times n$ ,  $n$  oznacza rozmiar sieci.



Przestrzeń rozwiązań:

$S = 2^N$ , czyli w przypadku sieci 10 x 10 komórek,  $S = 2^{100}$ .

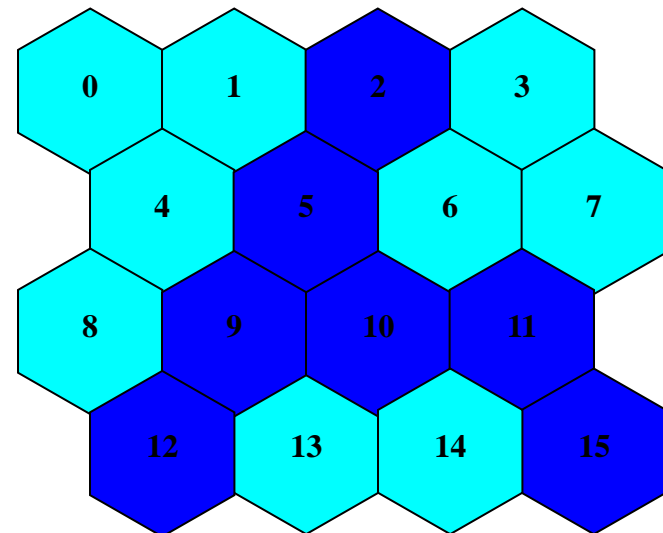
Przykład:

$n = 4$ ,  $N = 16$

$S = 2^{16} = 65536$



Chromosom algorytmu genetycznego



Rozkład stacji raportujących w sieci