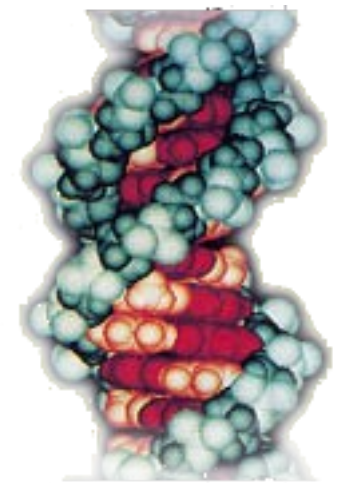
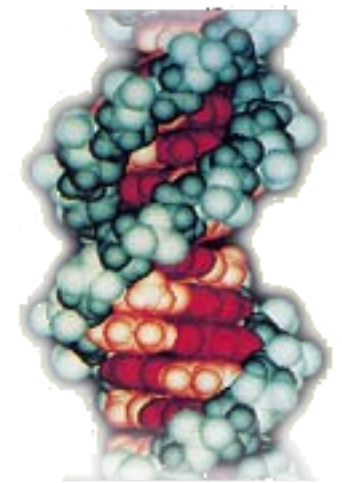


Algorytmy genetyczne i wielomiany w zagadnieniu interpolacji



Seminarium Metod Inteligencji Obliczeniowej
Warszawa 30 V 2007
mgr inż. Marcin Borkowski

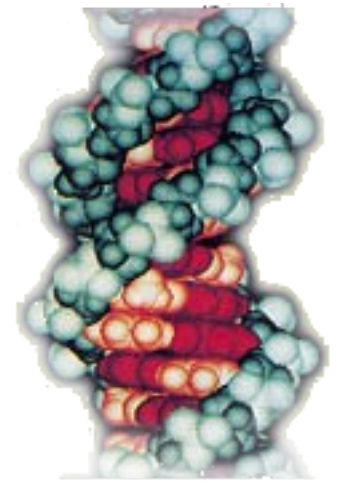
Dziś opowiem o:



- Algorytmie genetycznym i niszach
- Starszym wykorzystaniu tego algorytmu
- Szeregu logistycznym i wielomianach jako wzorcach
- Wynikach starszych, nowych i najnowszych
- A także trochę o najbliższych badaniach

Niszowy AG

Nabór



- Stały rozmiar populacji
- Kodowanie binarne
- Nabór:

Osobniki w populacji są sortowane względem ich przystosowania. Osobniki najslabsze są następnie zastępowane przez wyniki operatorów genetycznych.

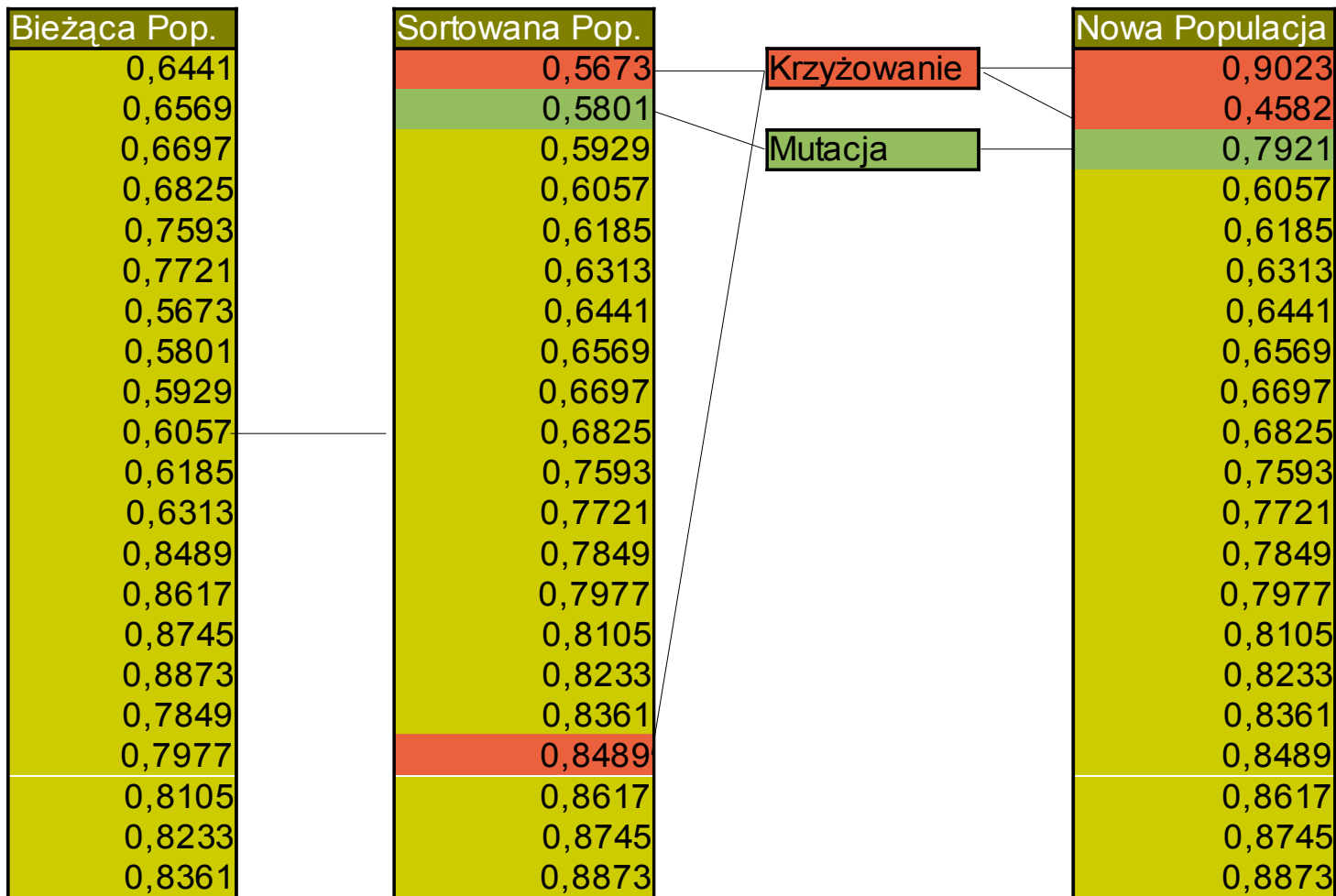
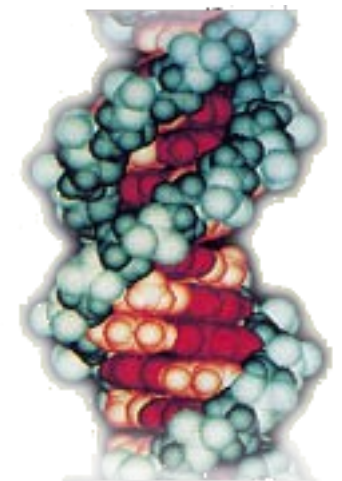
Pozwala to najlepszym osobnikom unikać “śmierci” tak długo aż zostaną wyparte przez osobniki jeszcze lepsze.

Wyniki operatorów są umieszczane w tymczasowej kopii populacji co daje równe szanse także osobnikom spisany na straty

Nie ma klonowania osobników

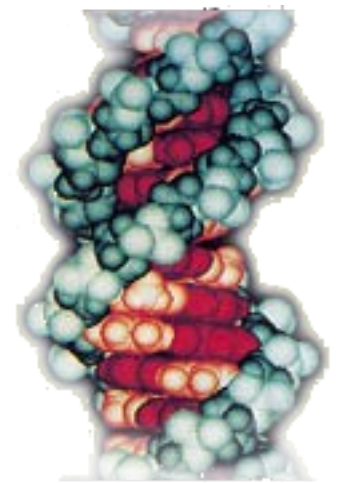
Niszowy AG

Nabór przykład



Niszowy AG

Mutacja



- Mutacja sterowana dwoma parametrami:

Każdy osobnik podlega mutacji z zadany
prawdopodobieństwem mutacji struktury pms .

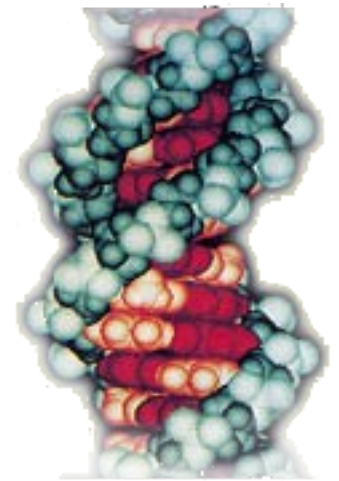
Każdy bit osobnika wylosowanego do mutacji podlega mutacji
(inwersja) z prawdopodobieństwem pm .

Pm maleje w kolejnych generacjach aż do granicy jednej mutacji
na osobnika na cykl

Wynik mutacji wypiera z populacji osobnika najslabszego z
jeszcze niepodmienionych

Niszowy AG

Krzyżowanie



- Krzyżowanie klasyczne jednopunktowe z prawdopodobieństwem pc

Z sumy wartości oczekiwanych pms i pc wynika że można oczekiwać, że najlepsze osobniki nie zostaną zniszczone

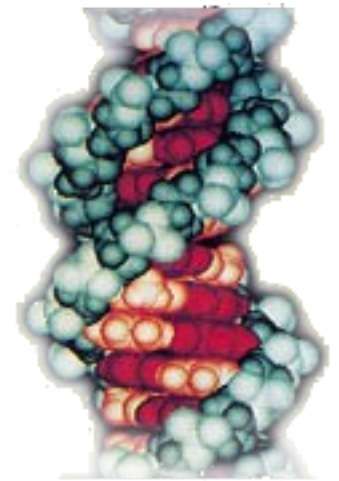
Nabór nie namnaża super osobników, dłuższe przetrwanie w populacji gwarantuje ich pozycja w rankingu

Kopie powstają rzadko na skutek znoszących się mutacji lub krzyżowania już istniejących kopii

Jeśli jednak już powstaną osobniki identyczne to nie będą one jednocześnie zajmować wysokiej pozycji w populacji – patrz ścisk

Niszowy AG

Ocena



- Funkcja oceny:

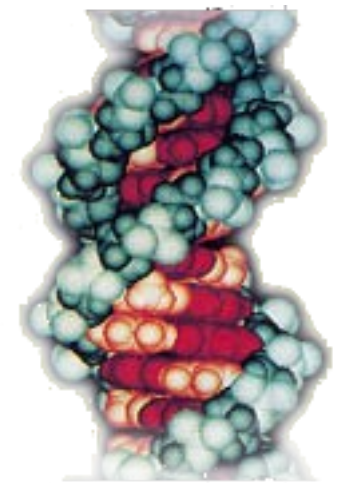
Ponieważ rozważane są głównie problemy dopasowania wzorca do serii danych konieczne jest policzenie wielu ocen – stopnia dopasowania dla wszystkich możliwych pozycji wzorca, do AG brana jest wartość uśredniona

Ocena rozbita jest na dwa etapy, poza zwyczajową ewaluacją osobnika następuje redukcja oceny na skutek działania czynnika ścisku – patrz ścisk

Ścisk wyznaczany jest dopiero gdy znane są wszystkie funkcje oceny z pierwszego etapu

Niszowy AG

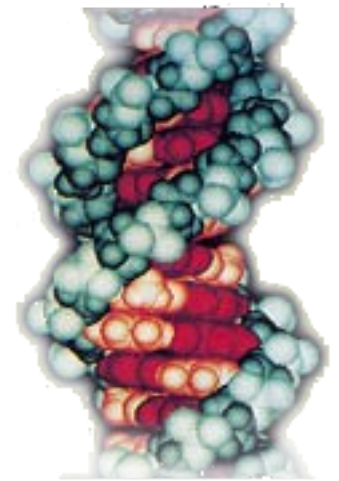
Nisze-ścisk



- Rozwiązań dostarczanych przez GA jest wiele, są to najlepsze optima w przeszukiwanej przestrzeni. Ilość tych optimów zależy od badanej przestrzeni (ilość optimów) i rozmiaru populacji
- Przy odpowiednio dobranym parametrze podobieństwa osobników jedno optimum przekłada się na jedną niszę
- Przeszukiwanie wielu optimów naraz jest możliwe dzięki ściskowi

Niszowy AG

Ścisk



W zamyśle ma działać jako narzędzie limitujące ilość podobnych do siebie osobników zamieszkujących tą samą niszę poprzez redukcowanie ich wartości dopasowania.

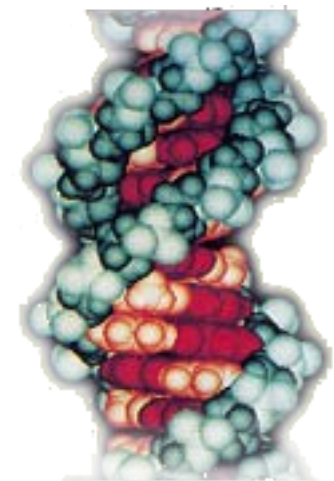
Nie wszystkie osobniki w niszy będą tak samo zredukowane - redukcja jest proporcjonalna do dotychczasowej wartości dopasowania.

Poza numeryczną wartością dopasowania pod uwagę brane są także inne czynniki np.: Długość wzorca, ilość wzorców uczących, przydatność itp.

Osobniki są podobne wtedy, gdy dla tych samych zadanych problemów dają podobne wyniki (podobna fizyczna budowa nie wystarcza)

Niszowy AG

Ścisk-przykład:



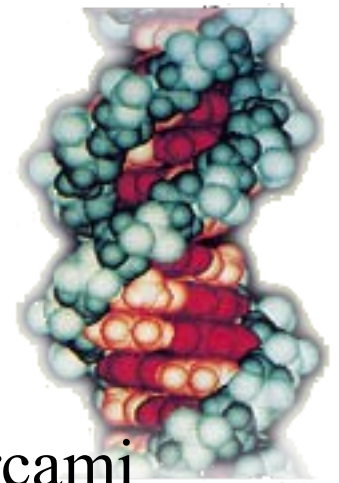
Osobniki A, B i C są podobne, tj. zajmują tą samą niszę

	Ocena	Długość wzorca	Suma ścisku	Ocena z ściskiem
<i>Osobnik A:</i>	<i>0,901</i>	<i>5</i>	<i>1</i>	<i>0,451</i>
<i>Osobnik B:</i>	<i>0,943</i>	<i>6</i>	<i>0</i>	<i>0,943</i>
<i>Osobnik C:</i>	<i>0,901</i>	<i>7</i>	<i>2</i>	<i>0,300</i>

Porównanie	A	B
Ścisk	1	0
Porównanie	A	C
Ścisk	0	1
Porównanie	B	C
Ścisk	0	1

Niszowy AG

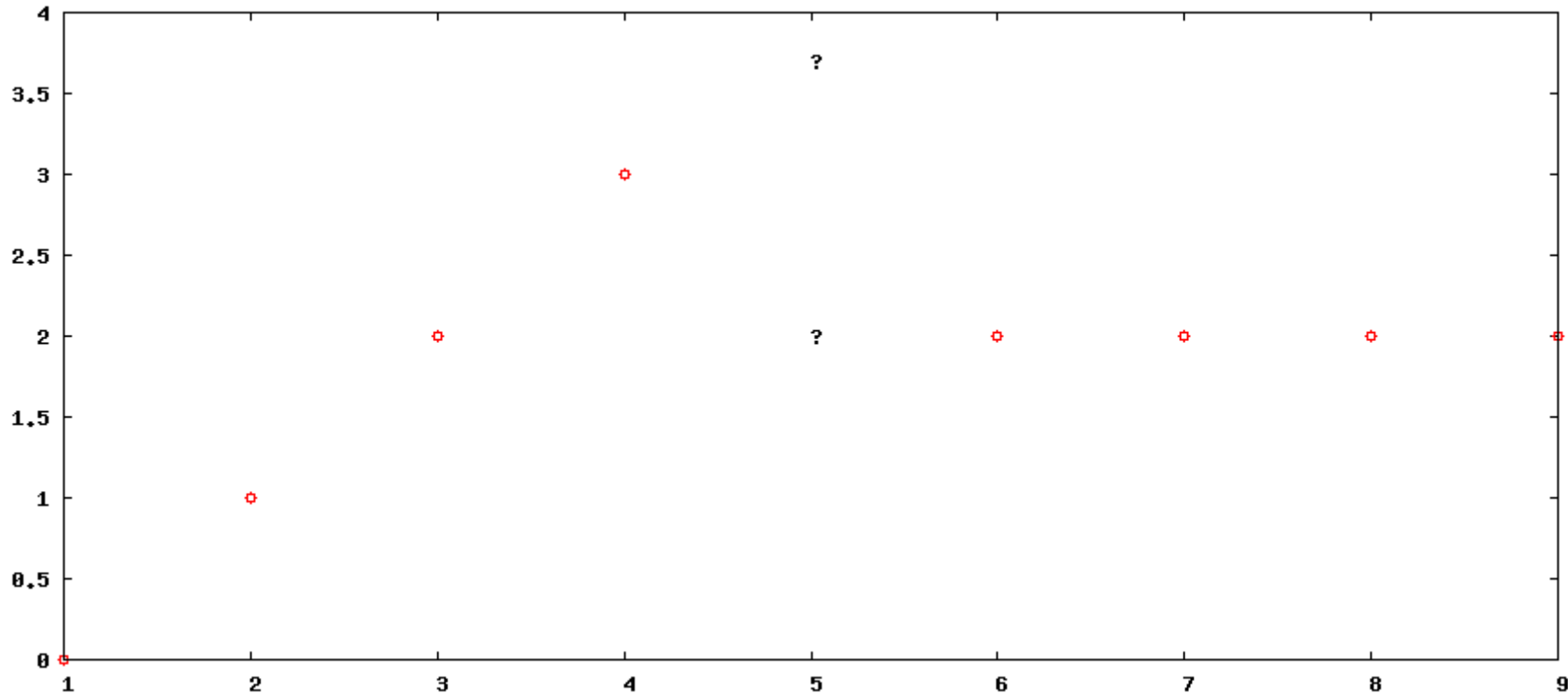
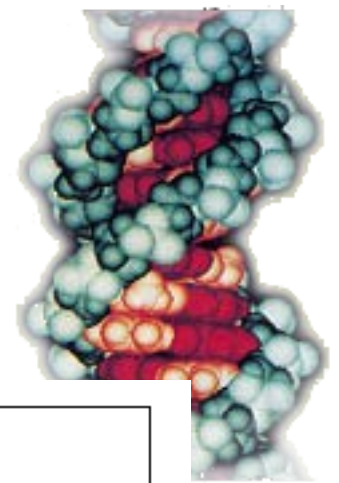
Ścisk



- Ścisk wspomaga rozłączne pokrycie serii danych wzorcami
- Osobniki należące do więcej niż jednej niszy “cierpią” na większy ścisk.
- Nisze mogą być związane z konkretnym odcinkiem serii danych, np, wzorce pasujące tylko na pewnym odcinku danych
- Osobniki pochodzące z różnych nisz mogą prowadzić do sprzecznych wyników na wspólnych częściach serii danych.
- Np.:

Niszowy AG

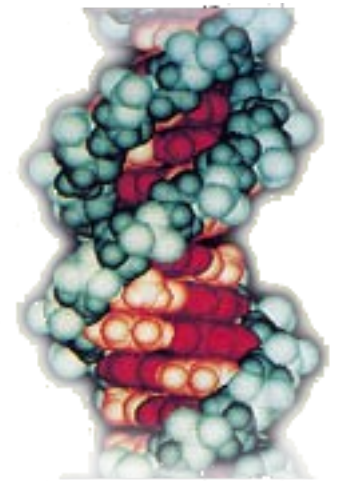
Ścisk



- Jaką wartość można interpolować dla pozycji 5 ?
- Rozwiązań może być wiele i mogą one być różne
- Niszowy AG znajdzie obie opcje (5,2) i (5,4)

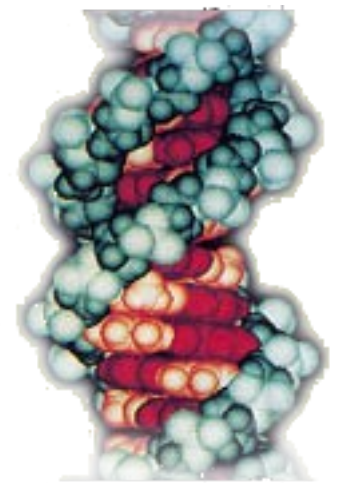
Niszowy AG

Uwagi



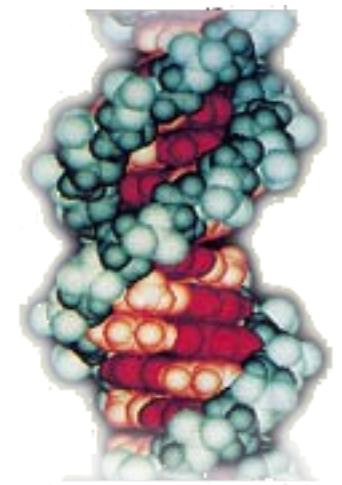
- Algorytm wychodzi od losowej populacji
- Po osiągnięciu pierwszych wyników zawsze można oczekiwać lepszych lub zmienić problem i pozwolić populacji prze-migrować do nowych nisz
- Nisze pozwalają szukać rozwiązań kompleksowych w jednym przebiegu AG, np.: wyszukać zbiór rozłącznych wzorców opisujących serię danych.
- Większość metodologii wyszukuje po jednym wzorcu na raz, a następnie dokonuje integracji rozwiązania, usuwa sprzeczności
- Niszowy AG nie wymaga integracji rozwiązań, ani usuwania sprzeczności

Proste zastosowanie



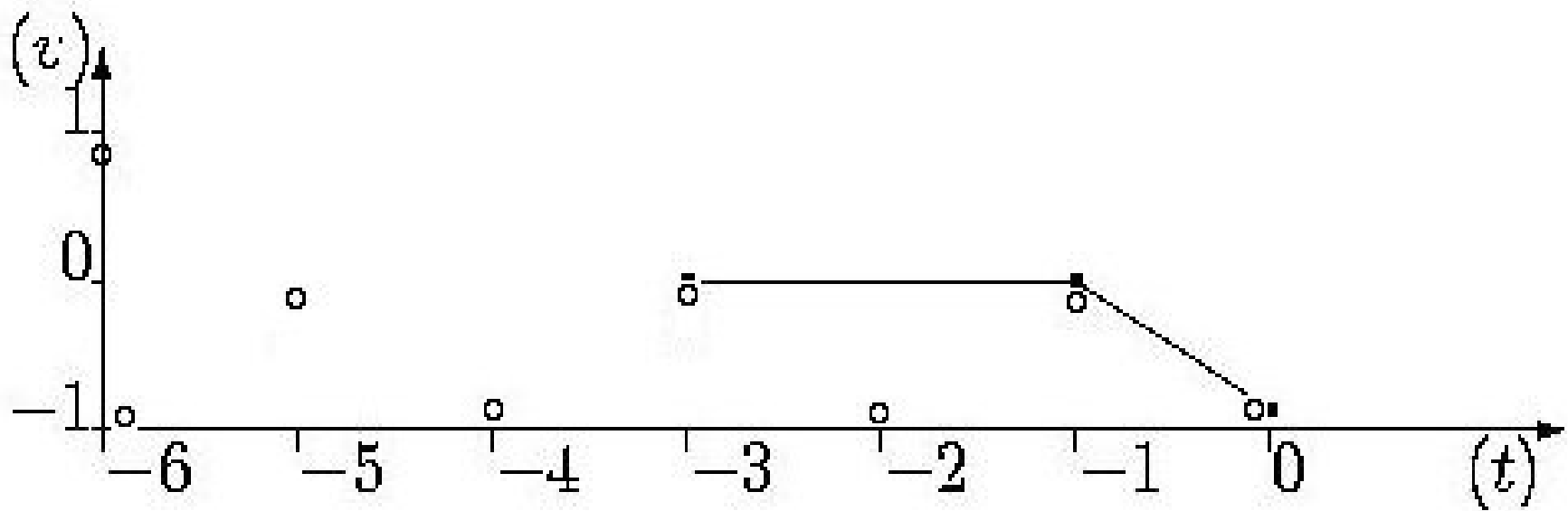
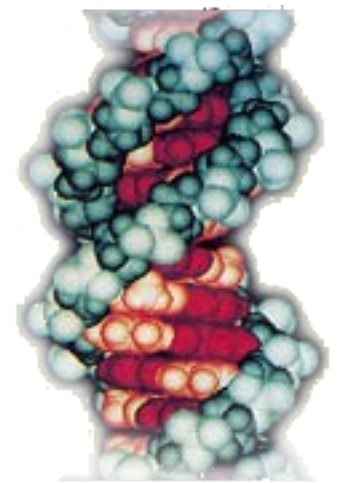
- Testowano:
 - Interpolacje
 - Predykcje
- Proste szeregi czasowe
 - Schodki
 - Sinus

Proste zastosowanie

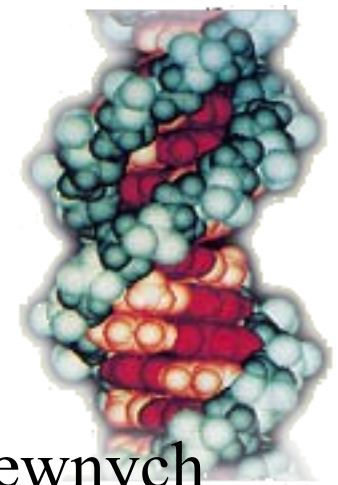


- Postać rozwiązania:
 - Wzorzec ma postać:
$$(v_1, 0)(v_2, 0_2) \dots (v_n, 0_n)$$
 - Odstęp pomiędzy węzłami jest limitowany
 - Dane wejściowe i wartości v_i są skalowane do przedziału $\langle -1, 1 \rangle$
 - Zastosowanie wzorca do predykcji lub interpolacji nie jest częścią algorytmu GA
 - Poszukiwanie wzorców można przeplatać z ich użyciem

Przykład wzorca

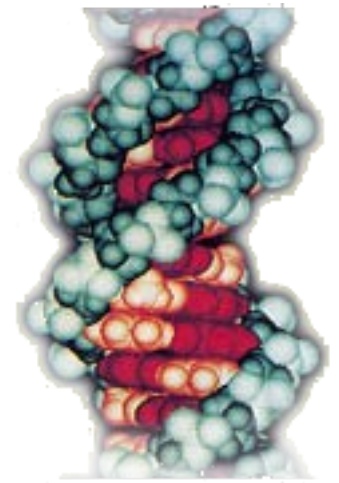


Proste zastosowanie



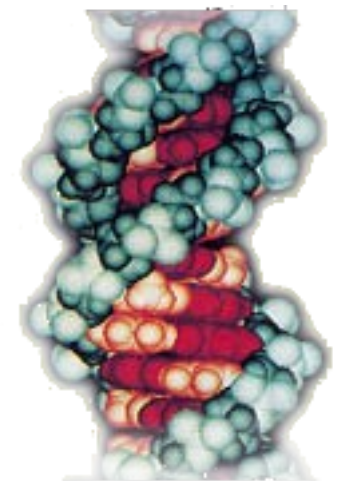
- Każdy wzorzec/osobnik sprawdza się tylko w pewnych określonych warunkach, stopień zgodności wzorca z konkretnymi danymi jest reprezentowany współczynnikiem zaufania
- Współczynnika zaufania jest proporcjonalny do odległości węzłów wzorca od rzeczywistych danych.
- Do zadania można włączyć liniowe skalowanie wzorców (wartości v_i) co rozszerza ich użyteczność
- Do rozwiązania są brane tylko wzorce o dużym współczynniku zaufania i wysokiej wartości oceny AG
- Możliwe jest kilka alternatywnych rozwiązań lub też kompletny brak rozwiązania

Proste zastosowanie



- Oceny dopasowania (fitness)
 - Do wyszukania wzorców wykorzystuje się dane uczące (znaną część szeregu).
 - Jedna znana dana jest z szeregu usuwana, a następnie uzupełniana przez wzorce
 - Jeśli predykcja jest zgodna z przewidywaniem - ocena cząstkowa jest wysoka, jeśli nie i jednocześnie współczynnik zaufania jest wysoki – ocena cząstkowa jest niska
 - Proces ten jest powtarzany dla wszystkich znanych danych w szeregu i wszystkich osobników-wzorców
 - Finalna ocena jest uśredniana z najlepszych (pod względem wsp. zaufania) ocen cząstkowych

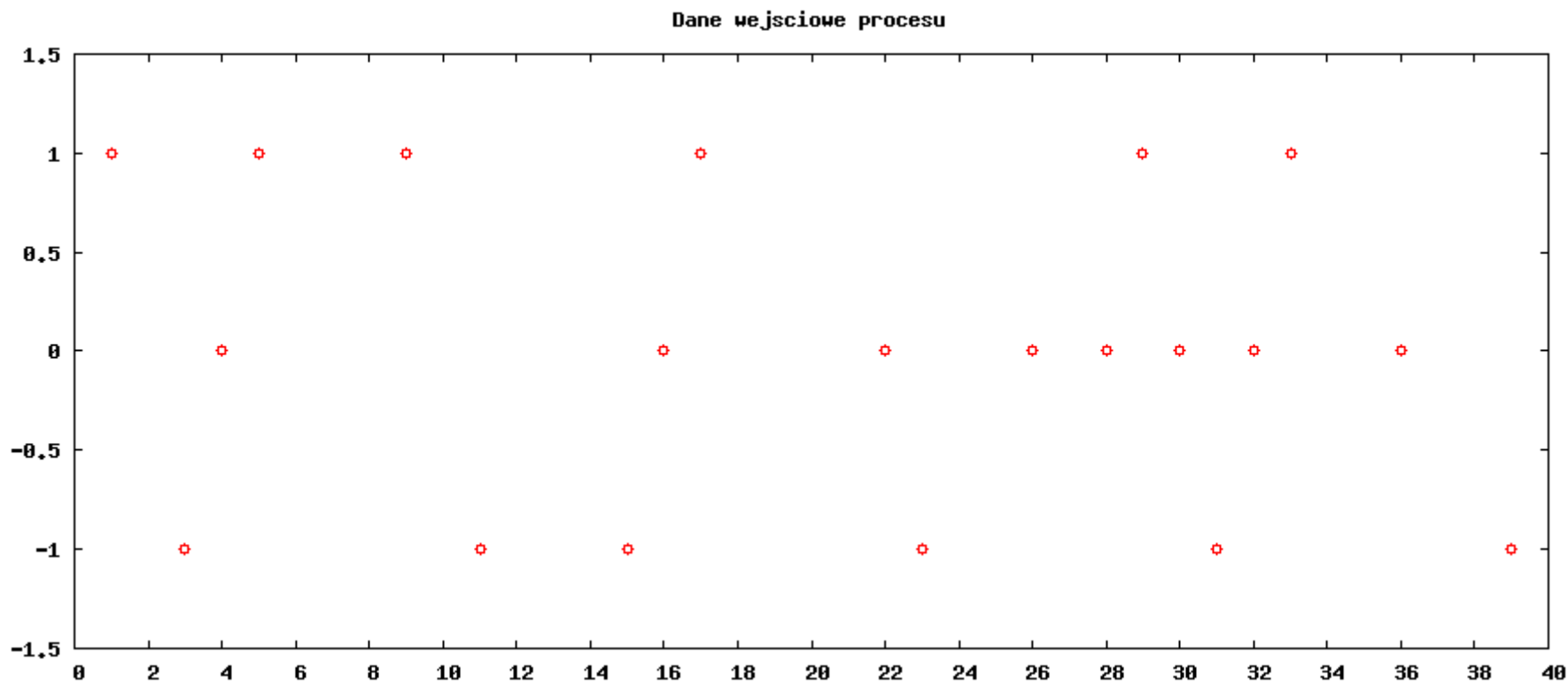
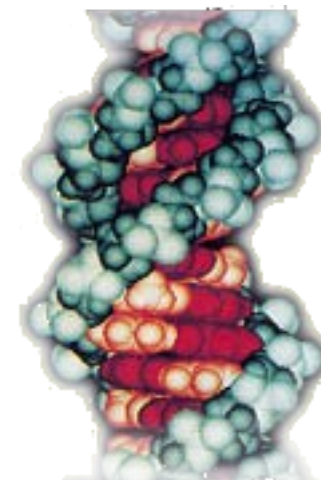
Wyniki – dla prostych wzorców



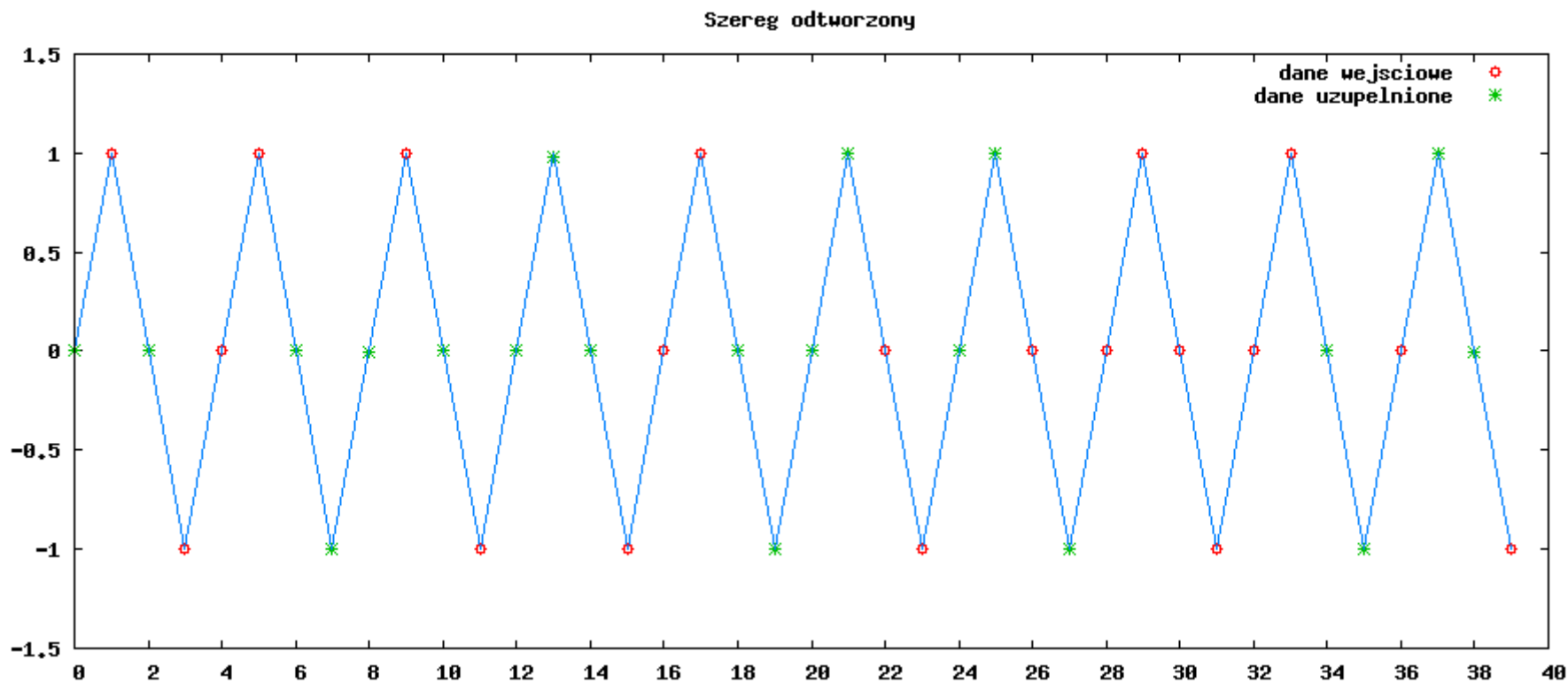
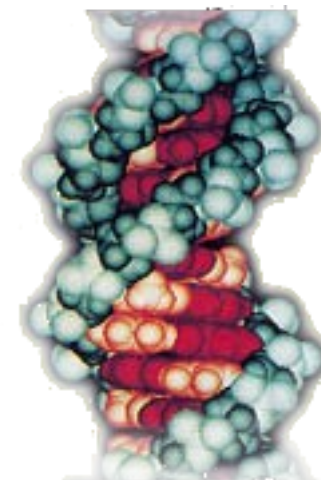
Parametry procesu

- ilość bitów wartości w węźle 12
- maksymalny dystans między węzłami 4
- ilość węzłów we wzorcu $\langle 3,4 \rangle$
- prawdopodobieństwo krzyżowania 45%
- prawdopodobieństwo mutacji struktury 50%
- rozmiar populacji 70
- Początkowe generacje AG $\langle 300,3000 \rangle$
- Wtórne kroki AG $\langle 100,300 \rangle$
- Oczekiwany minimalny poziom zaufania $\langle 75,95 \rangle$

Wyniki – dla prostych wzorców



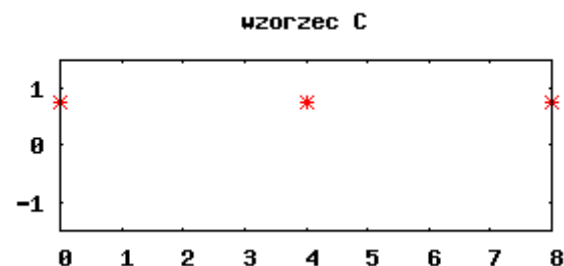
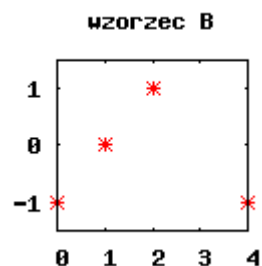
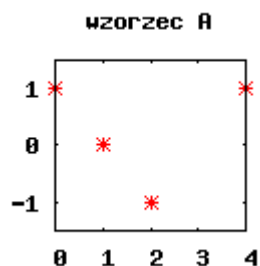
Wyniki – dla prostych wzorców



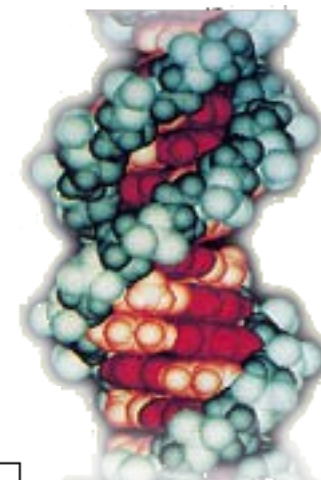
Wyniki – dla prostych wzorców



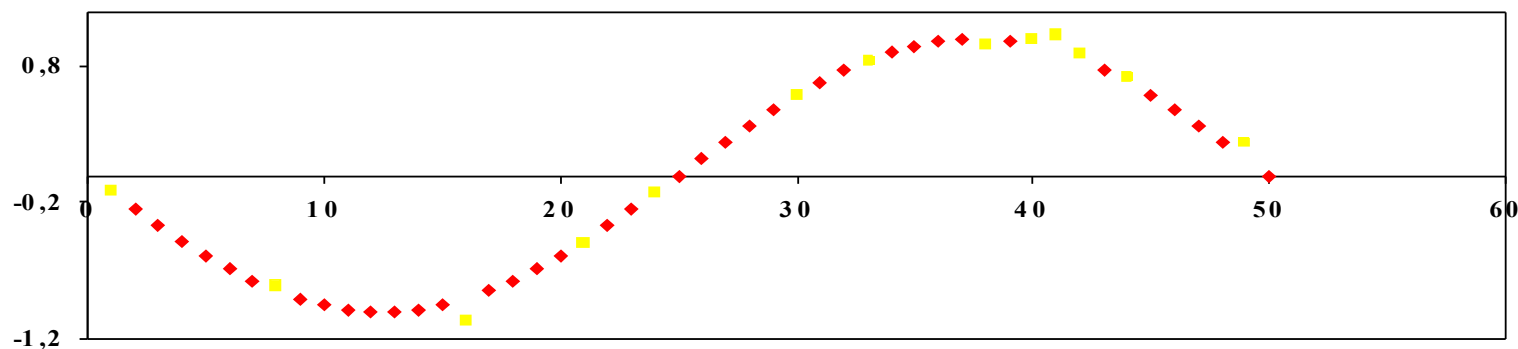
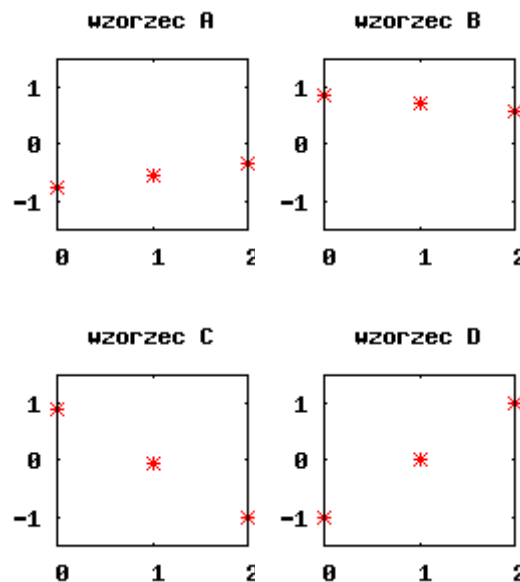
- Odtwarzanie prostego szeregu (52% braków - kwadraty)
- Najlepsze wzorce bez skalowania:
A) $(1.00,0)(0.00,1)(-1.00,2)(1.00,4)$
B) $(-1.00,0)(0.00,1)(1.00,2)(-1.00,4)$
- Najlepsze wzorce ze skalowaniem:
C) $(0.768,0)(0.768,4)(0.768,8)$



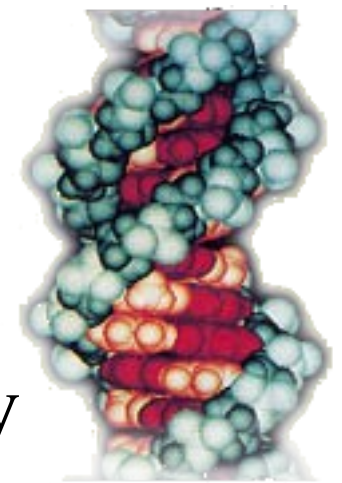
Wyniki – dla prostych wzorców



- Sinus – 26 % braków
- Najlepsze wzorce ze skalowaniem:
A) $(-0.75, 0)(-0.55, 1)(-0.35, 2)$
B) $(0.86, 0)(0.72, 1)(0.59, 2)$
C) $(0.89, 0)(-0.07, 1)(-0.99, 2)$
D) $(-1.00, 0)(0.03, 1)(0.99, 2)$

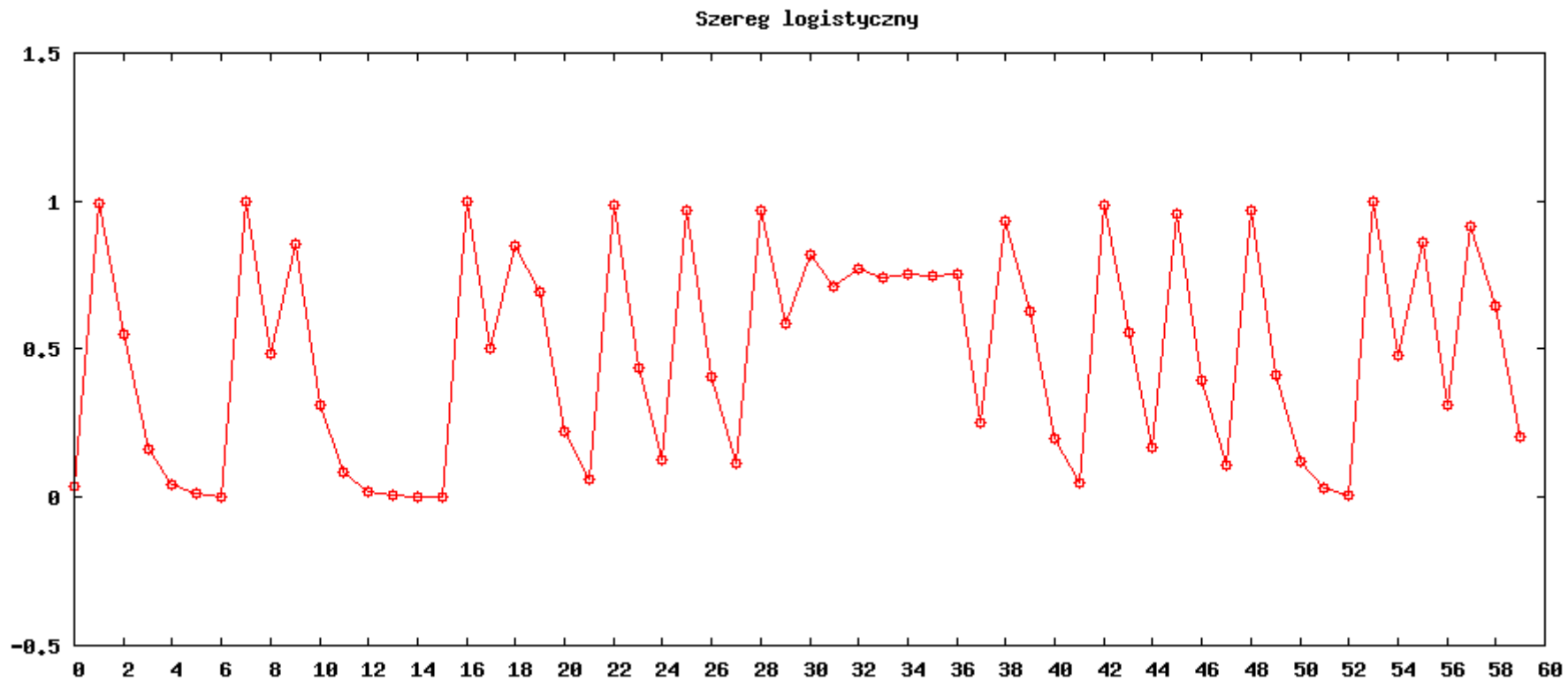
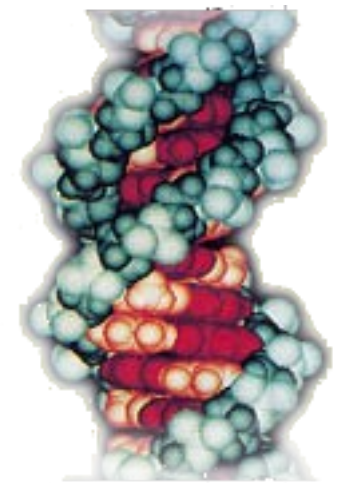


Szereg logistyczny

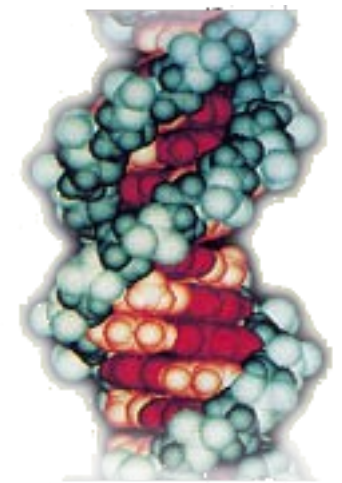


- Równanie logistyczne – system chaotyczny
- $X_{n+1} = R * X_n (1 - X_n)$
- R – stałe
- Wyrazy utrzymują się w przedziale $[0,1]$ dla:
 - X_0 w przedziale $[0,1]$
 - R w przedziale $(0,4]$
 - Dla R z przedziału $[3.57,4]$ rozkład wartości zbliża się do jednostajnego, a zachowanie funkcji jest chaotyczne
- Użyty szereg $x_0 = 0.2027$ $R = 4$

Szereg w postaci używanej w testach

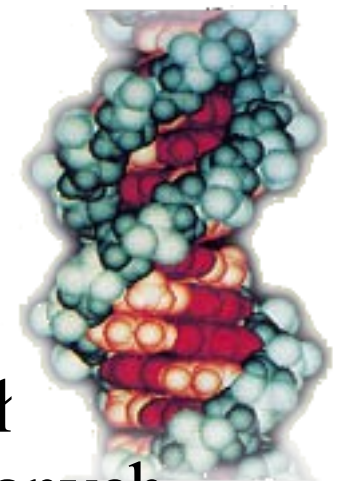


Wielomiany jako wzorce



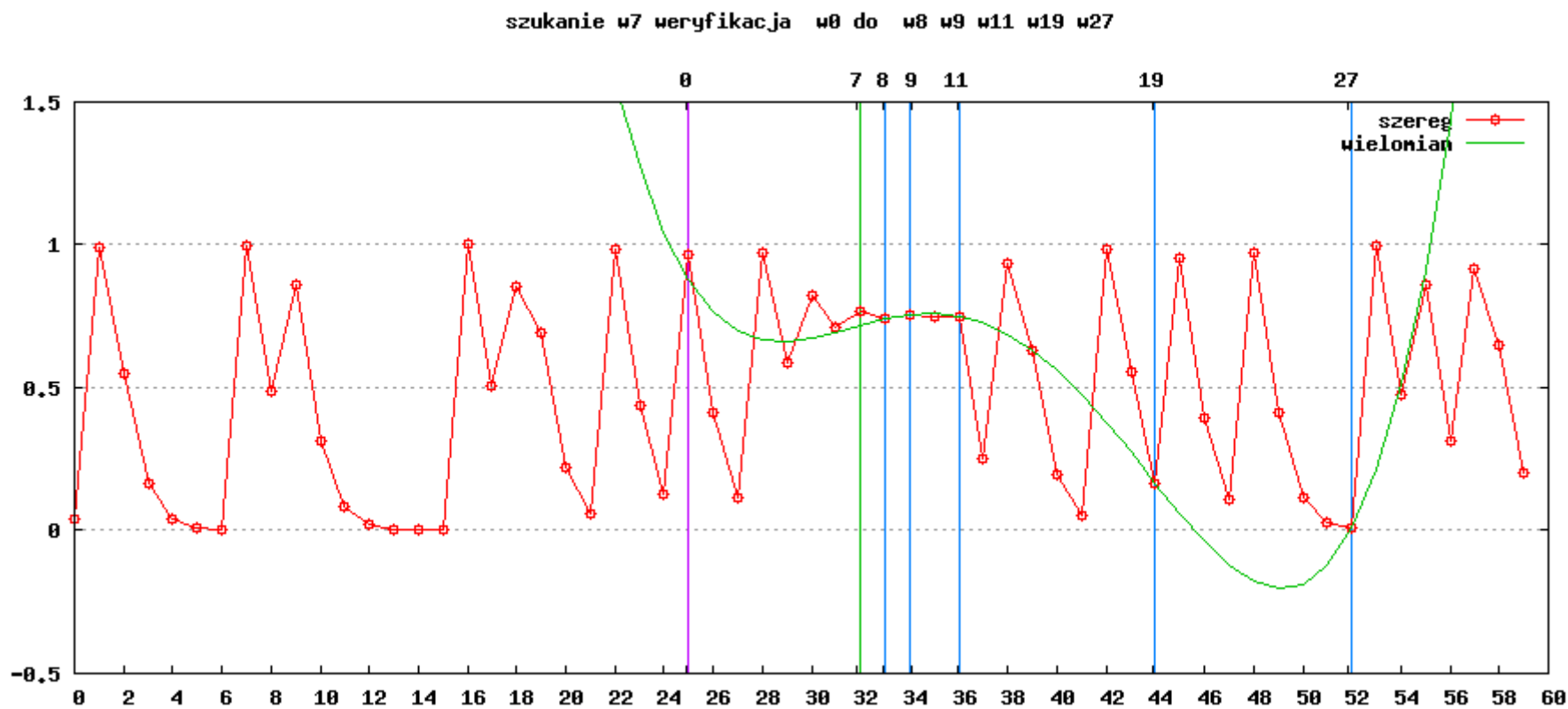
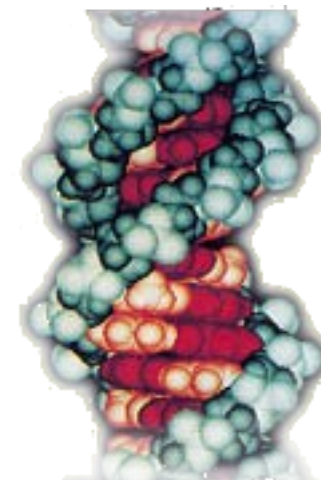
- Wielomiany n-tego rzędu:
 - $a_n X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0$
 - Konstrukcja wielomianu na podstawie znanych wartości, konieczne $n+1$ węzłów
 - Wartości w węzłach są zawsze brane z istniejących danych, więc AG ma za zadanie jedynie wyszukać rozkład węzłów
 - Weryfikacja wyniku - dodatkowy $n+2$ węzeł
 - Wynik generowane przez AG to lista przesunięć węzłów względem siebie np.: 0 7 8 9 11 19 27

Wielomiany jako wzorce

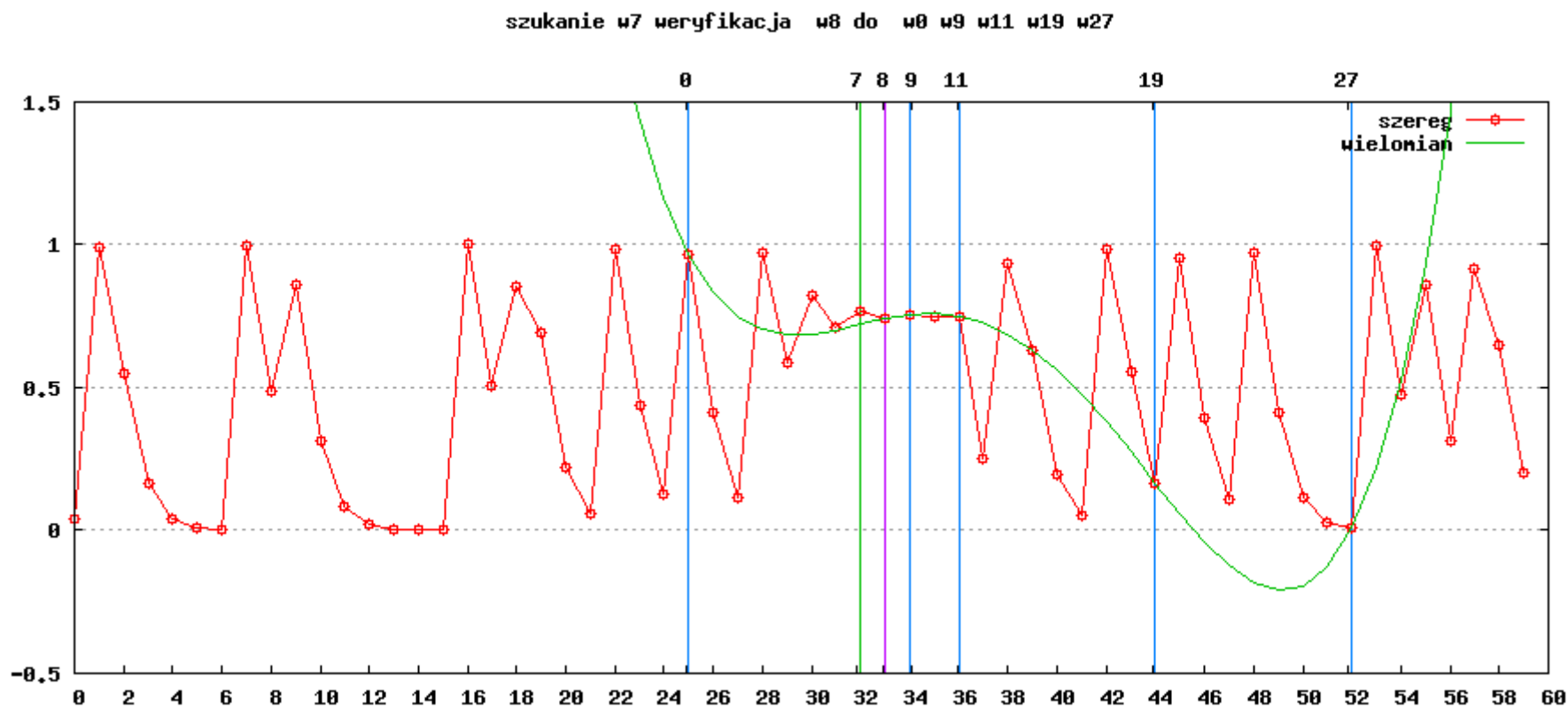
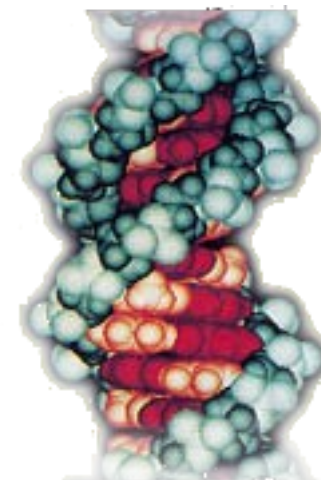


- W tego rozkładu (0 7 8 9 11 19 27) jeden węzeł (1 z $n+3$) będzie ustawiał wzorzec względem danych,
- W praktyce sprawdza się wszystkie $n+3$ (w przykładzie 7) takie ustawienia i wybiera to o najwyższym współczynniku zaufania
- Dodatkowo każdy z węzłów już znanych bierze udział w wyznaczaniu współczynnika zaufania czyli określaniu stopnia w jakim rozkład pasuje do danych.
- Przykład:

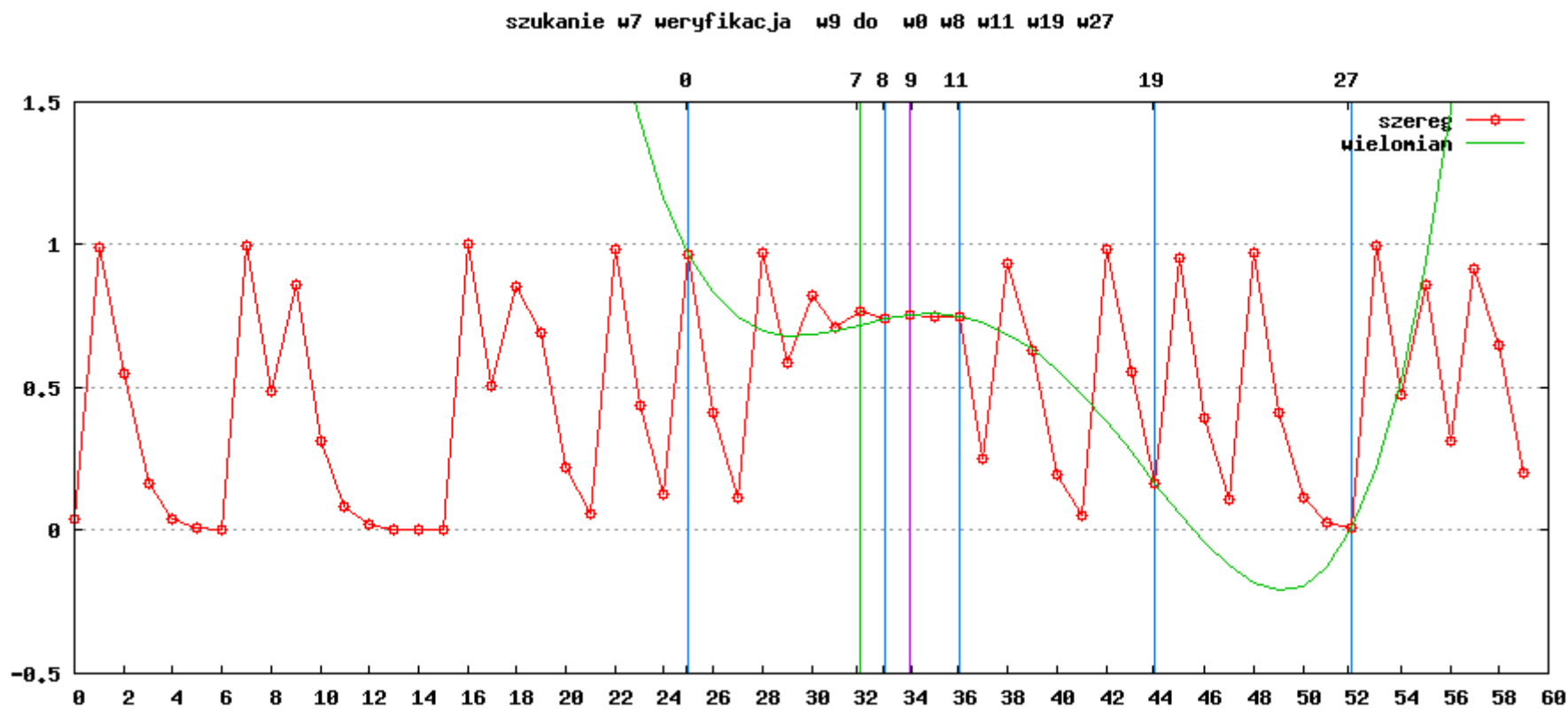
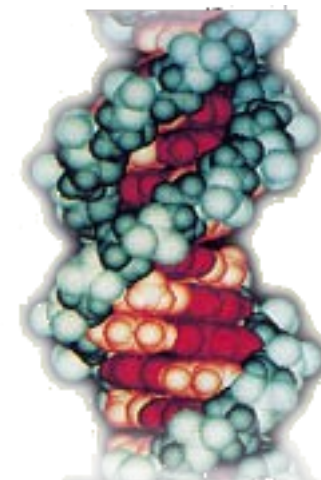
Wielomiany jako wzorce



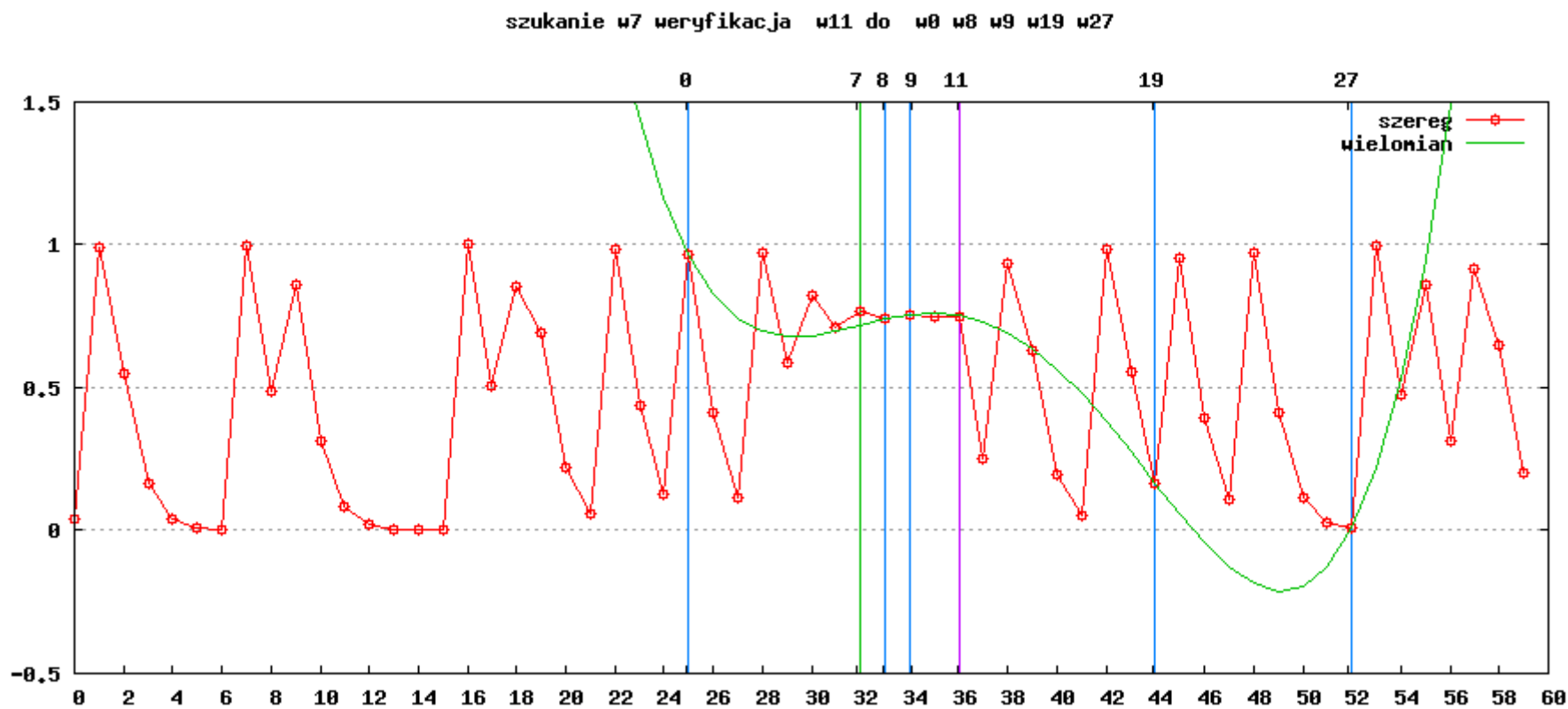
Wielomiany jako wzorce



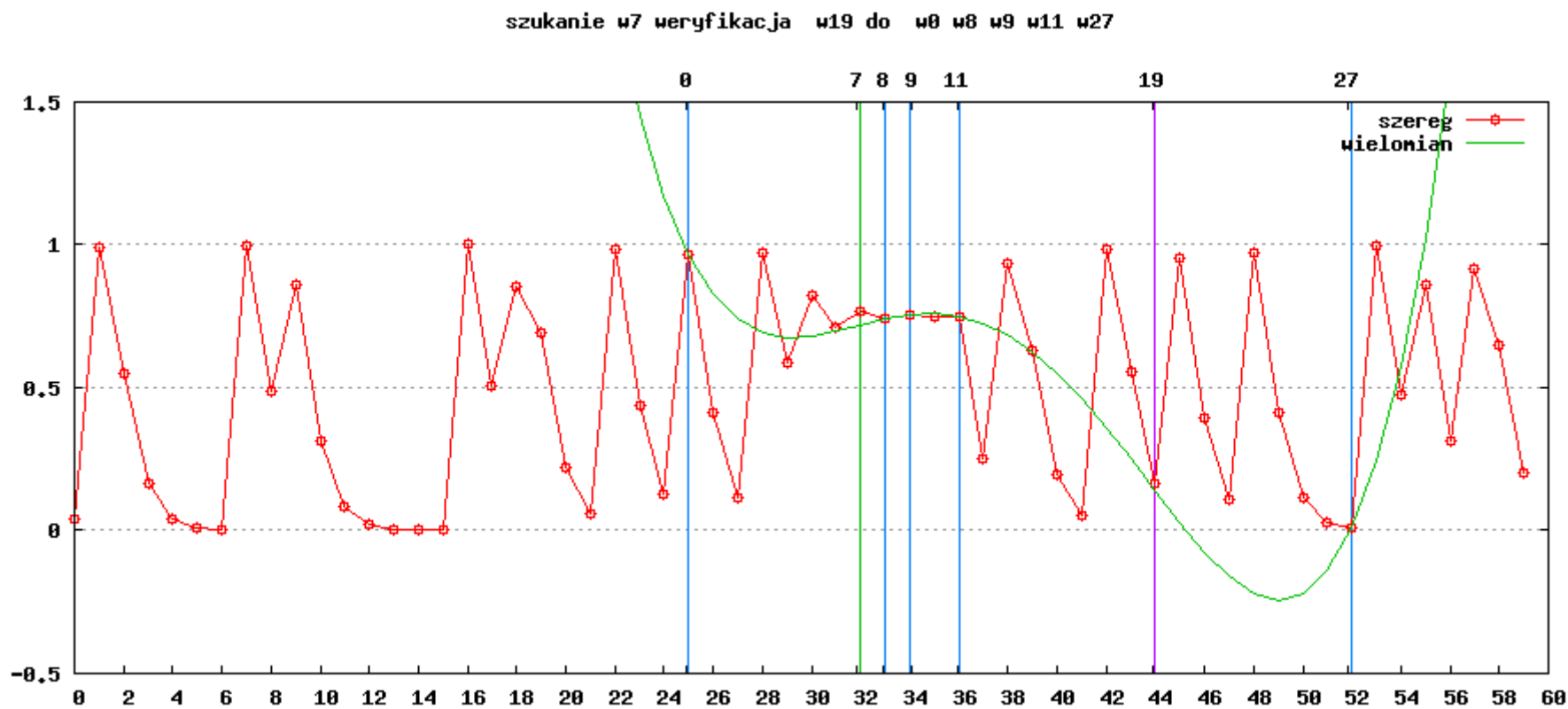
Wielomiany jako wzorce



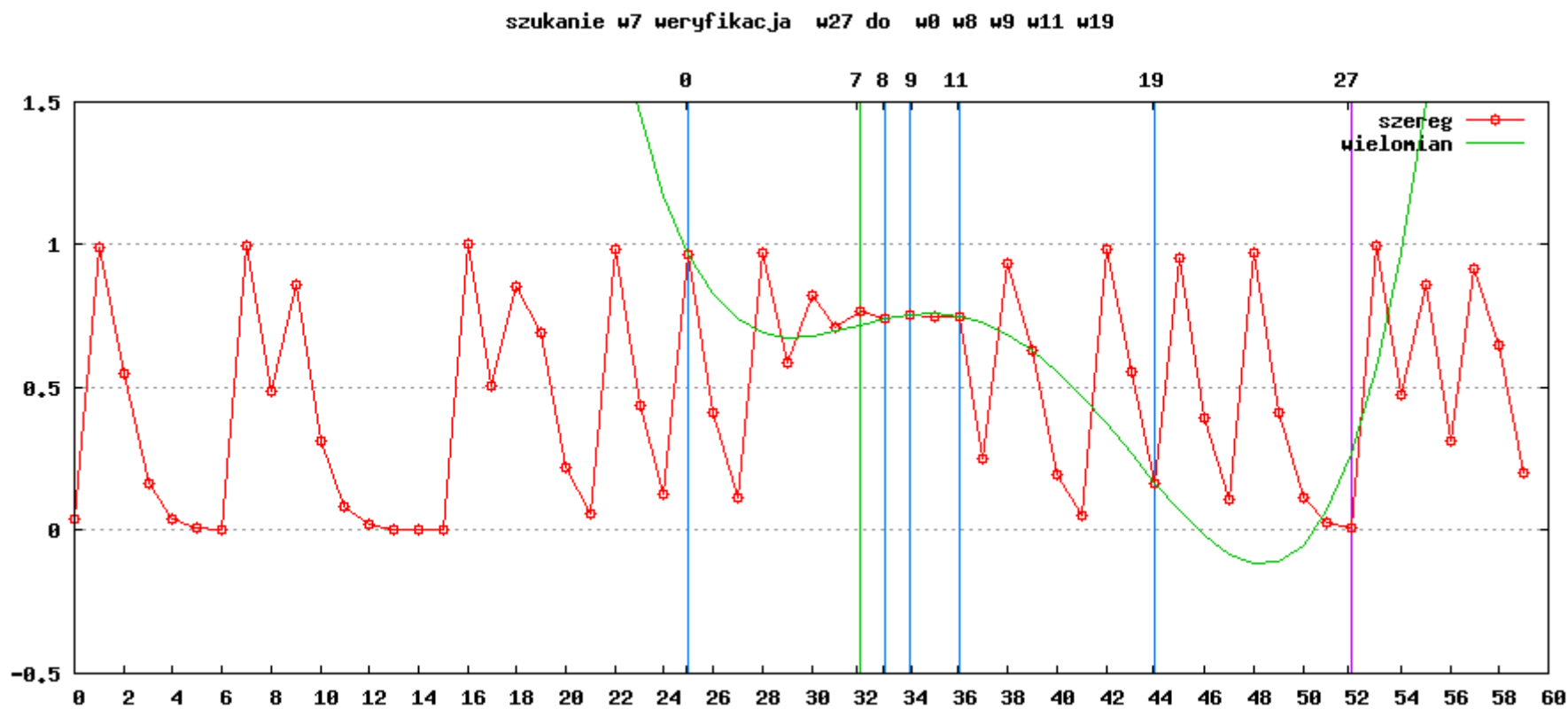
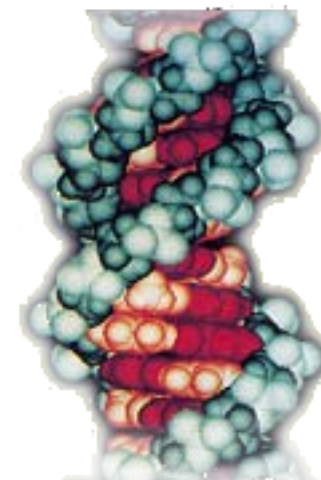
Wielomiany jako wzorce



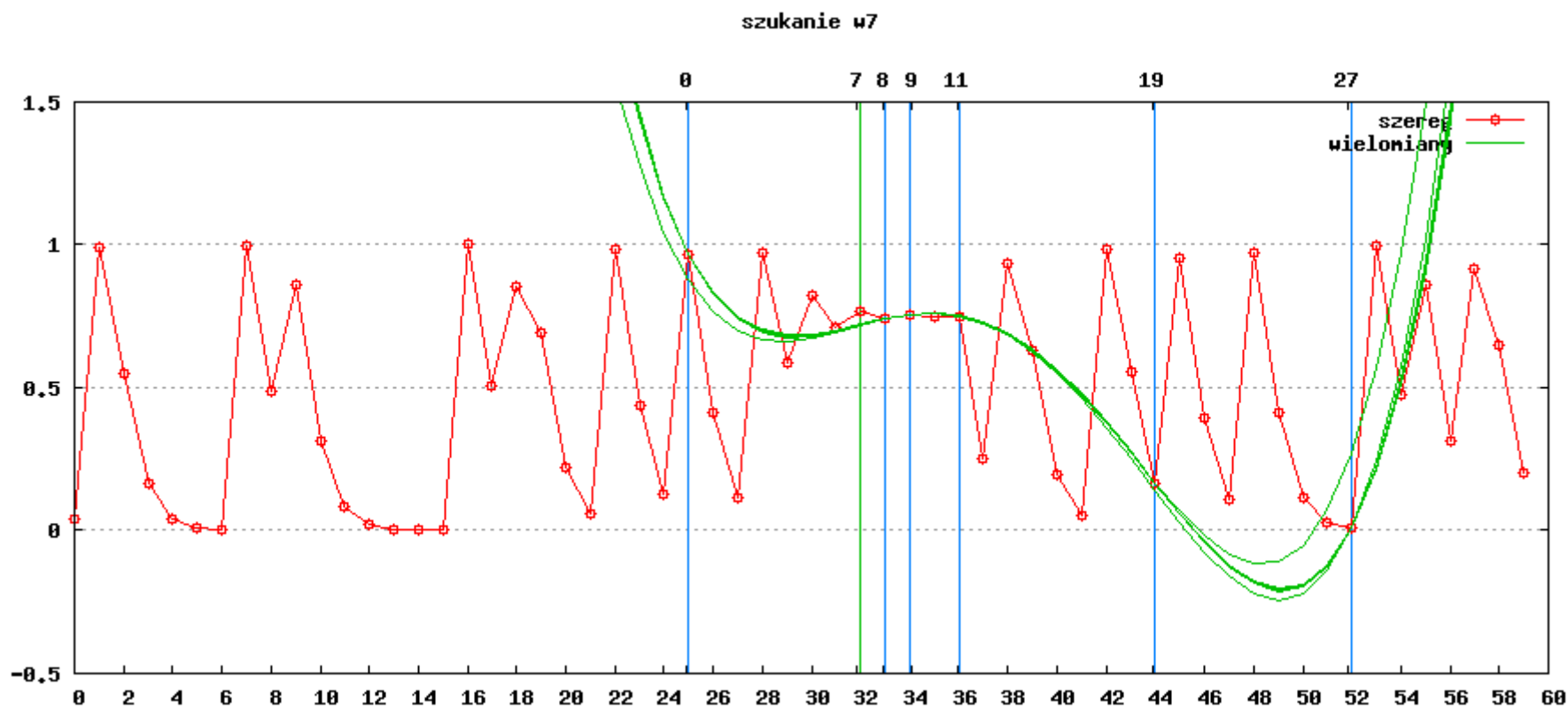
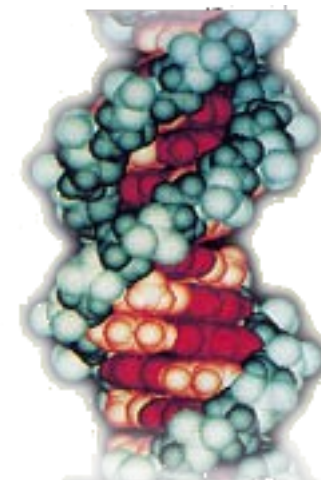
Wielomiany jako wzorce



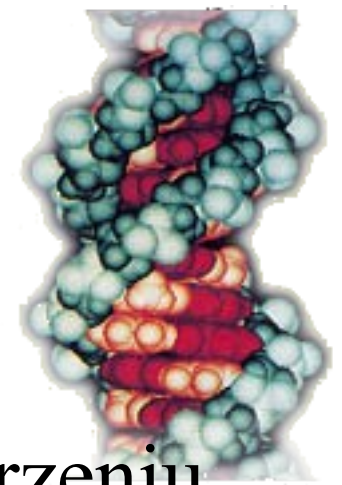
Wielomiany jako wzorce



Wielomiany jako wzorce



Wielomiany jako wzorce



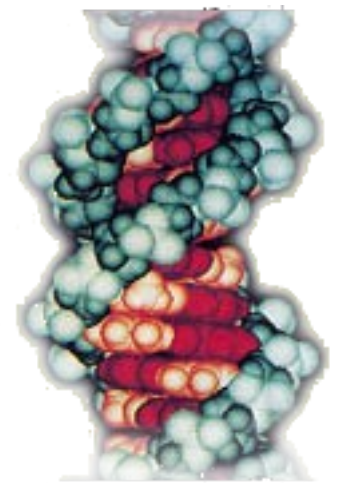
- Z $(n+2)$ węzłów rozkładu służących stworzeniu wielomianu wybierane są wszystkie kombinacje $(n+1)$ węzłów, numeryczna, uśredniona różnica na pozostawionym węźle stanowi współczynnik zaufania.
- Przy okazji $n+2$ razy wyznaczamy też brakujący węzeł, odchylenie standardowe w tym zbiorze wartości powinno być małe i wpływa na wiarygodność wzorca.

Testy



- Z 60 wyrazowego szeregu logistycznego usuwamy jeden element z przedziału [20,39]
- Wykonujemy 5 testów z 1 brakiem
- Powtarzamy test dla wszystkich 20 wartości usuwanych
- Powtarzamy cały test dla wielomianów od stopnia 4 do 8
- Razem $20 * 5 * 5 = 500$ testów

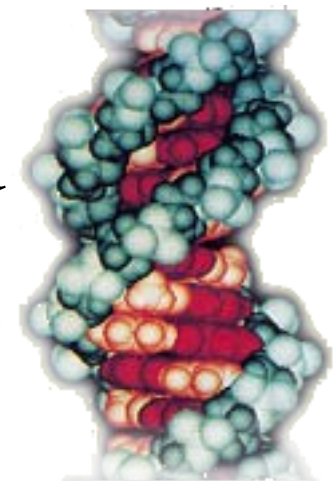
Parametry testów



Parametry procesu

- maksymalny dystans między węzłami 8
- ilość węzłów we wzorcu 7-11
- prawdopodobieństwo krzyżowania 35%
- prawdopodobieństwo mutacji struktury 50%
- rozmiar populacji 100
- Początkowe generacje 2000
- Wtórne kroki 0
- Oczekiwany minimalny poziom zaufania 78%

Pojedynczy raport z testu



```
p=0.313(91.3%|0.771afit|2c),c 1 at 39 node 4 TP=0 rkg=0.063208 tCheat=11 tTP=9
p=0.801(97.7%|0.789afit|2c),c 1 at 39 node 3 TP=0 rkg=0.071515 tCheat=14 tTP=14
p=0.100(83.4%|0.832afit|0c),c 1 at 39 node 3 TP=0 rkg=0.648868 tCheat=6 tTP=6
p=0.882(95.0%|0.775afit|0c),c 1 at 39 node 5 TP=0 rkg=0.701799 tCheat=6 tTP=5*
```

...

```
-----POPULATION:          2001-----
```

```
(0) (8) (10) (11) (13) (14) (15) (18) [TP=19,C=0,U=0] [f=0.425/2.878=0.148,age= 1,m0]
```

```
(0) (2) ( 8) (14) (17) (20) (28) (36) [TP= 5,C=7,U=1] [f=0.748/2.942=0.254,age= 1,m0]
```

...

```
(0) (8) (10) (12) (14) (17) (19) (23) [TP= 2,C=1,U=0] [f=0.909/1.000=0.909,age= 24,m1]
```

```
(0) (7) (11) (15) (23) (29) (31) (37) [TP= 1,C=1,U=1] [f=0.911/1.000=0.911,age= 303,m0]
```

```
(0) (6) ( 7) (10) (11) (12) (14) (17) [TP= 1,C=0,U=0] [f=0.917/1.000=0.917,age= 127,m0]
```

```
(0) (8) (13) (14) (15) (21) (22) (28) [TP= 1,C=0,U=0] [f=0.922/1.000=0.922,age= 78,m2]
```

```
(0) (7) (11) (14) (16) (17) (18) (19) [TP= 1,C=0,U=0] [f=0.930/1.000=0.930,age= 520,m2]
```

```
(0) (2) ( 4) ( 8) (11) (12) (20) (27) [TP= 1,C=2,U=0] [f=0.930/1.000=0.930,age= 883,m0]
```

```
(0) (4) ( 5) ( 8) (10) (11) (13) (21) [TP= 1,C=2,U=0] [f=0.934/1.000=0.934,age= 34,m2]
```

```
(0) (7) ( 9) (10) (14) (20) (22) (30) [TP= 1,C=0,U=0] [f=0.950/1.000=0.950,age=1218,m0]
```

```
(0) (2) ( 8) (10) (13) (14) (22) (30) [TP= 8,C=2,U=0] [f=0.957/1.000=0.957,age= 960,m0]
```

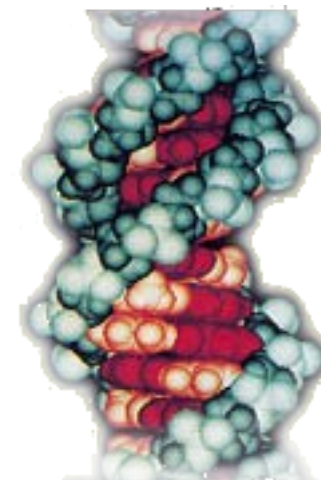
```
(0) (7) (11) (12) (14) (15) (17) (18) [TP= 1,C=0,U=1] [f=0.958/1.000=0.958,age= 520,m0]
```

```
(0) (7) ( 8) (10) (11) (13) (17) (25) [TP= 1,C=1,U=0] [f=0.960/1.000=0.960,age= 45,m0]
```

```
(0) (4) ( 5) ( 6) ( 8) ( 9) (10) (17) [TP= 1,C=1,U=0] [f=0.973/1.000=0.973,age= 40,m0]
```

```
-----
```

Wyniki dopasowań procentowo:



Rząd wielomianu - 4

0.223	20	:	15.20%	10.59%	27.40%	36.38%	7.70%	srd:19.45	naj.:07.70
0.059	21	:	4.68%	22.71%	22.71%	38.88%	4.68%	srd:18.73	naj.:04.68
0.985	22	:	41.10%	41.10%	41.10%	41.10%	41.10%	srd:41.10	naj.:41.10
0.439	23	:	2.98%	2.98%	2.98%	2.98%	13.88%	srd: 5.16	naj.:02.98
0.125	24	:	16.69%	5.97%	3.16%	5.97%	30.02%	srd:12.36	naj.:03.16
0.968	25	:	24.92%	12.25%	41.25%	7.04%	29.92%	srd:23.07	naj.:07.04
0.410	26	:	0.61%	0.61%	6.35%	0.59%	8.03%	srd: 3.23	naj.:00.59
0.116	27	:	4.89%	4.89%	4.89%	7.54%	17.33%	srd: 7.90	naj.:04.89
0.970	28	:	15.67%	22.32%	15.67%	22.32%	22.32%	srd:19.66	naj.:15.67
0.586	29	:	16.34%	0.90%	0.90%	0.90%	27.59%	srd: 9.32	naj.:00.90
0.822	30	:	1.89%	8.11%	7.90%	1.89%	7.67%	srd: 5.49	naj.:01.89
0.711	31	:	2.68%	4.84%	4.84%	4.84%	0.34%	srd: 3.50	naj.:00.34
0.769	32	:	2.84%	5.39%	2.40%	1.87%	20.07%	srd: 6.51	naj.:01.87
0.740	33	:	4.20%	0.55%	1.78%	2.40%	2.40%	srd: 2.26	naj.:00.55
0.755	34	:	1.71%	1.71%	1.71%	0.33%	1.04%	srd: 1.30	naj.:00.33
0.748	35	:	1.89%	9.56%	0.17%	3.59%	0.58%	srd: 3.15	naj.:00.17
0.751	36	:	3.91%	7.62%	31.87%	31.87%	31.87%	srd:21.42	naj.:03.91
0.251	37	:	22.72%	25.83%	25.32%	33.34%	28.75%	srd:27.19	naj.:22.72
0.933	38	:	10.66%	22.75%	22.75%	22.68%	22.75%	srd:20.31	naj.:10.66
0.630	39	:	0.30%	11.73%	5.73%	12.05%	16.50%	srd: 9.26	naj.:00.30

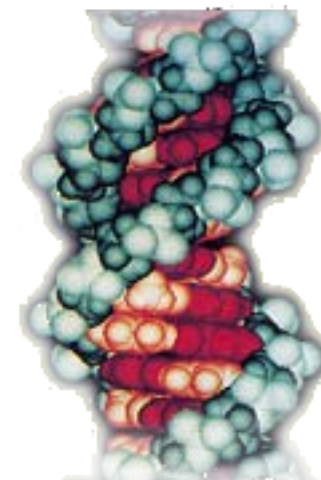
Wyniki dopasowań procentowo:



Rząd wielomianu - 5

0.223	20	:	27.05%	15.37%	30.62%	19.63%	27.05%	srd:23.94	naj.:15.37
0.059	21	:	41.55%	12.77%	6.67%	15.84%	44.19%	srd:24.20	naj.:06.67
0.985	22	:	20.42%	10.07%	30.00%	25.17%	47.20%	srd:26.57	naj.:10.07
0.439	23	:	12.20%	23.92%	27.20%	12.95%	16.06%	srd:18.46	naj.:12.20
0.125	24	:	19.91%	34.99%	13.84%	13.84%	11.03%	srd:18.72	naj.:11.03
0.968	25	:	28.00%	46.74%	5.96%	6.79%	28.64%	srd:23.22	naj.:05.96
0.410	26	:	18.35%	1.60%	19.56%	19.51%	5.45%	srd:12.89	naj.:01.60
0.116	27	:	18.08%	18.08%	18.08%	26.50%	26.50%	srd:21.44	naj.:18.08
0.970	28	:	21.95%	5.72%	35.56%	15.98%	4.01%	srd:16.64	naj.:04.01
0.586	29	:	2.32%	2.32%	3.95%	28.30%	2.62%	srd: 7.90	naj.:02.32
0.822	30	:	8.55%	3.95%	7.30%	9.16%	8.85%	srd: 7.56	naj.:03.95
0.711	31	:	2.66%	3.82%	0.95%	0.95%	14.05%	srd: 4.48	naj.:00.95
0.769	32	:	1.93%	3.28%	2.52%	0.49%	8.10%	srd: 3.26	naj.:00.49
0.740	33	:	0.39%	5.21%	1.89%	1.57%	12.17%	srd: 4.24	naj.:00.39
0.755	34	:	4.40%	0.45%	3.23%	5.48%	1.38%	srd: 2.98	naj.:00.45
0.748	35	:	1.00%	21.08%	1.00%	5.90%	2.95%	srd: 6.38	naj.:01.00
0.751	36	:	0.70%	2.26%	10.77%	0.87%	6.11%	srd: 4.14	naj.:00.70
0.251	37	:	19.73%	25.50%	27.24%	25.39%	20.37%	srd:23.64	naj.:19.73
0.933	38	:	38.29%	32.34%	12.59%	21.73%	12.59%	srd:23.50	naj.:12.59
0.630	39	:	12.63%	21.02%	26.28%	18.15%	19.78%	srd:19.57	naj.:12.63

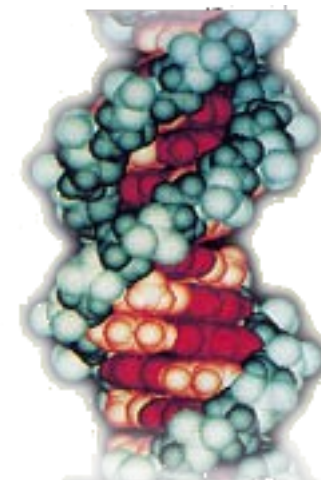
Wyniki dopasowań procentowo:



Rząd wielomianu - 6

0.223	20	:	8.55%	4.35%	0.30%	22.80%	24.37%	srd:12.07	naj.:00.30
0.059	21	:	28.70%	39.46%	43.41%	11.52%	12.45%	srd:27.10	naj.:11.52
0.985	22	:	27.33%	44.93%	31.02%	28.14%	34.54%	srd:33.19	naj.:27.33
0.439	23	:	17.28%	10.08%	25.96%	10.85%	0.76%	srd:12.98	naj.:00.76
0.125	24	:	1.59%	38.59%	12.49%	7.46%	0.11%	srd:12.04	naj.:00.11
0.968	25	:	35.18%	36.41%	6.83%	36.41%	11.77%	srd:25.32	naj.:06.83
0.410	26	:	14.60%	15.81%	7.52%	16.47%	17.12%	srd:14.30	naj.:07.52
0.116	27	:	37.68%	28.25%	26.42%	15.73%	8.15%	srd:23.24	naj.:08.15
0.970	28	:	11.48%	12.61%	21.46%	20.34%	13.85%	srd:15.94	naj.:11.48
0.586	29	:	16.31%	0.77%	14.13%	12.55%	19.44%	srd:12.64	naj.:00.77
0.822	30	:	9.02%	8.76%	7.93%	4.60%	15.88%	srd: 9.23	naj.:04.60
0.711	31	:	2.58%	2.08%	2.38%	3.53%	4.76%	srd: 3.06	naj.:02.08
0.769	32	:	6.89%	3.21%	2.44%	2.87%	9.13%	srd: 4.90	naj.:02.44
0.740	33	:	2.38%	1.62%	2.30%	7.57%	2.21%	srd: 3.21	naj.:01.62
0.755	34	:	0.48%	2.65%	5.34%	23.10%	0.21%	srd: 6.35	naj.:00.21
0.748	35	:	1.91%	7.60%	10.87%	8.17%	1.28%	srd: 5.96	naj.:01.28
0.751	36	:	4.54%	20.38%	0.49%	9.54%	7.58%	srd: 8.50	naj.:00.49
0.251	37	:	34.25%	7.93%	24.22%	12.85%	8.63%	srd:17.57	naj.:07.93
0.933	38	:	9.75%	27.25%	0.14%	14.88%	38.05%	srd:18.01	naj.:00.14
0.630	39	:	31.35%	17.17%	13.16%	12.45%	13.05%	srd:17.43	naj.:12.45

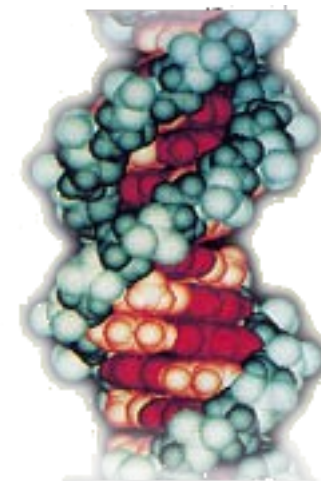
Wyniki dopasowań procentowo:



Rząd wielomianu - 7

0.223	20	:	22.98%	24.28%	3.51%	28.49%	12.69%	srd:18.39	naj.:03.51
0.059	21	:	0.15%	34.22%	5.90%	36.85%	4.42%	srd:16.30	naj.:00.15
0.985	22	:	9.23%	28.41%	49.16%	46.18%	46.33%	srd:35.86	naj.:09.23
0.439	23	:	6.38%	0.25%	16.83%	20.30%	2.21%	srd: 9.19	naj.:00.25
0.125	24	:	14.09%	42.66%	10.23%	1.60%	32.08%	srd:20.13	naj.:01.60
0.968	25	:	5.82%	4.14%	1.74%	9.77%	38.90%	srd:12.07	naj.:01.74
0.410	26	:	11.68%	27.35%	27.25%	19.90%	21.29%	srd:21.49	naj.:11.68
0.116	27	:	29.30%	33.34%	40.43%	20.33%	18.06%	srd:28.29	naj.:18.06
0.970	28	:	13.78%	33.75%	14.68%	18.67%	6.56%	srd:17.48	naj.:06.56
0.586	29	:	9.24%	2.05%	9.47%	5.45%	22.49%	srd: 9.74	naj.:02.05
0.822	30	:	8.38%	5.86%	6.10%	7.94%	1.92%	srd: 6.04	naj.:01.92
0.711	31	:	4.40%	2.60%	2.50%	1.74%	0.96%	srd: 2.44	naj.:00.96
0.769	32	:	5.27%	0.29%	6.44%	19.80%	0.89%	srd: 6.53	naj.:00.29
0.740	33	:	2.90%	24.08%	2.37%	3.28%	6.78%	srd: 7.88	naj.:02.37
0.755	34	:	1.45%	4.16%	4.59%	0.92%	5.96%	srd: 3.41	naj.:00.92
0.748	35	:	26.42%	1.51%	6.84%	6.10%	4.15%	srd: 9.00	naj.:01.51
0.751	36	:	11.09%	21.50%	0.92%	5.13%	17.96%	srd:11.32	naj.:00.92
0.251	37	:	13.11%	13.68%	23.46%	22.32%	24.65%	srd:19.44	naj.:13.11
0.933	38	:	1.64%	42.89%	5.33%	24.83%	0.37%	srd:15.01	naj.:00.37
0.630	39	:	13.81%	22.62%	12.36%	10.10%	4.25%	srd:12.62	naj.:04.25

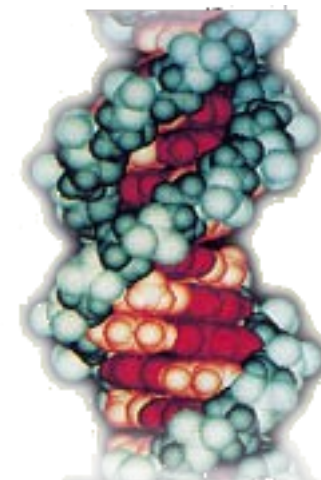
Wyniki dopasowań procentowo:



Rząd wielomianu - 8

0.223	20	:	12.71%	20.84%	26.21%	16.07%	10.28%	srd:17.22	naj.:10.28
0.059	21	:	9.04%	30.82%	24.65%	21.79%	23.06%	srd:21.87	naj.:09.04
0.985	22	:	44.32%	39.15%	29.44%	46.42%	0.46%	srd:31.95	naj.:00.46
0.439	23	:	6.24%	25.82%	21.91%	1.87%	20.90%	srd:15.34	naj.:01.87
0.125	24	:	32.09%	38.54%	7.54%	37.92%	33.86%	srd:29.99	naj.:07.54
0.968	25	:	7.74%	47.00%	17.47%	14.55%	32.14%	srd:23.78	naj.:07.74
0.410	26	:	20.75%	3.83%	17.29%	1.60%	17.90%	srd:12.27	naj.:01.60
0.116	27	:	25.30%	4.52%	22.55%	22.70%	35.96%	srd:22.20	naj.:04.52
0.970	28	:	34.23%	21.07%	13.96%	34.20%	23.52%	srd:25.39	naj.:13.96
0.586	29	:	23.15%	27.81%	2.45%	7.92%	22.38%	srd:16.74	naj.:02.45
0.822	30	:	6.94%	2.17%	6.87%	7.38%	6.07%	srd: 5.88	naj.:02.17
0.711	31	:	1.24%	4.28%	10.75%	23.73%	17.60%	srd:11.52	naj.:01.24
0.769	32	:	21.51%	26.37%	0.03%	8.86%	6.98%	srd:12.75	naj.:00.03
0.740	33	:	27.41%	7.05%	10.53%	1.55%	10.20%	srd:11.34	naj.:01.55
0.755	34	:	0.52%	3.59%	33.46%	8.96%	1.04%	srd: 9.51	naj.:00.52
0.748	35	:	4.34%	6.77%	6.93%	3.34%	4.99%	srd: 5.27	naj.:03.34
0.751	36	:	13.60%	7.29%	3.45%	3.64%	3.08%	srd: 6.21	naj.:03.08
0.251	37	:	15.76%	24.24%	28.29%	23.62%	16.99%	srd:21.78	naj.:15.76
0.933	38	:	2.54%	1.31%	5.53%	37.54%	14.48%	srd:12.28	naj.:01.31
0.630	39	:	14.81%	18.70%	18.50%	8.02%	11.94%	srd:14.39	naj.:08.02

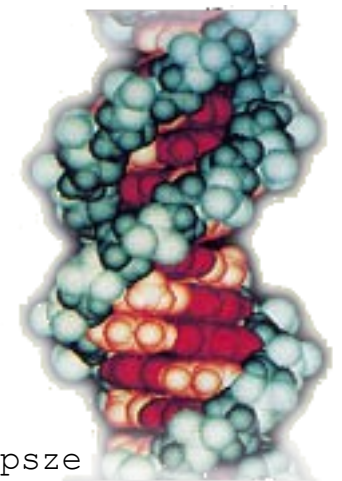
Wyniki dopasowań ranking:



Rząd wielomianu - 8

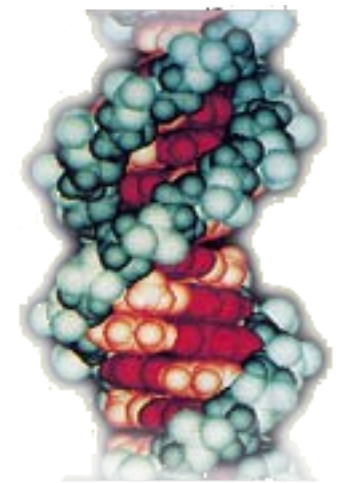
0.223	20:	12.71%(0.724444)	20.84%(0.765938)	26.21%(0.837773)	16.07%(0.790409)	10.28%(0.885074)
0.059	21:	9.04%(0.807686)	30.82%(0.773029)	24.65%(0.596818)	21.79%(0.066362)	23.06%(0.710447)
0.985	22:	44.32%(0.848788)	39.15%(0.747470)	29.44%(0.732933)	46.42%(0.748516)	0.46%(0.693359)
0.439	23:	6.24%(0.068456)	25.82%(0.622926)	21.91%(0.740005)	1.87%(0.660347)	20.90%(0.698620)
0.125	24:	32.09%(0.060455)	38.54%(0.719247)	7.54%(0.834346)	37.92%(0.762985)	33.86%(0.771402)
0.968	25:	7.74%(0.802390)	47.00%(0.868875)	17.47%(0.649744)	14.55%(0.792647)	32.14%(0.781297)
0.410	26:	20.75%(0.750966)	3.83%(0.594811)	17.29%(0.798752)	1.60%(0.754922)	17.90%(0.658508)
0.116	27:	25.30%(0.805419)	4.52%(0.871498)	22.55%(0.605148)	22.70%(0.774333)	35.96%(0.736398)
0.970	28:	34.23%(0.716201)	21.07%(0.847875)	13.96%(0.722056)	34.20%(0.639378)	23.52%(0.767464)
0.586	29:	23.15%(0.725734)	27.81%(0.059743)	2.45%(0.820305)	7.92%(0.072372)	22.38%(0.818804)
0.822	30:	6.94%(0.784925)	2.17%(0.863097)	6.87%(0.532114)	7.38%(0.720357)	6.07%(0.844175)
0.711	31:	1.24%(0.773711)	4.28%(0.833771)	10.75%(0.800532)	23.73%(0.717039)	17.60%(0.773480)
0.769	32:	21.51%(0.764407)	26.37%(0.891236)	0.03%(0.643422)	8.86%(0.880487)	6.98%(0.640107)
0.740	33:	27.41%(0.742536)	7.05%(0.696842)	10.53%(0.803344)	1.55%(0.795428)	10.20%(0.700366)
0.755	34:	0.52%(0.777574)	3.59%(0.763076)	33.46%(0.818747)	8.96%(0.743977)	1.04%(0.801834)
0.748	35:	4.34%(0.645959)	6.77%(0.870159)	6.93%(0.714118)	3.34%(0.679215)	4.99%(0.877048)
0.751	36:	13.60%(0.802187)	7.29%(0.059538)	3.45%(0.737092)	3.64%(0.771423)	3.08%(0.701118)
0.251	37:	15.76%(0.790839)	24.24%(0.807144)	28.29%(0.742835)	23.62%(0.772972)	16.99%(0.703652)
0.933	38:	2.54%(0.793647)	1.31%(0.727241)	5.53%(0.740973)	37.54%(0.720305)	14.48%(0.721122)
0.630	39:	14.81%(0.724455)	18.70%(0.825387)	18.50%(0.678275)	8.02%(0.828525)	11.94%(0.794215)

Automatyczny wybór najlepszych:



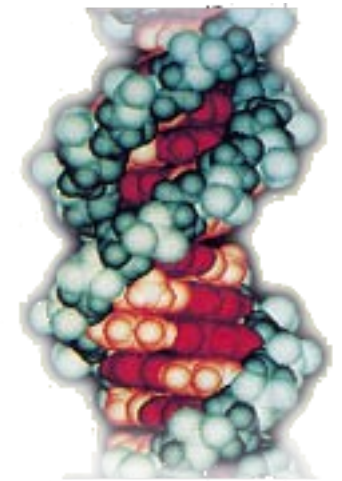
Rząd		4	5	6	7	8	Najlepsze
0.223	20:	07.70 (0.718)	15.37 (0.812)	00.30 (0.909)	03.51 (0.864)	10.28 (0.885)	00.30 (0.909)
0.059	21:	04.68 (0.719)	06.67 (0.725)	11.52 (0.746)	00.15 (0.866)	09.04 (0.807)	00.15 (0.866)
0.985	22:	41.10 (0.801)	47.20 (0.876)	27.33 (0.850)	09.23 (0.795)	00.46 (0.848)	47.20 (0.876)
0.439	23:	02.98 (0.752)	12.95 (0.881)	00.76 (0.810)	02.21 (0.903)	01.87 (0.740)	02.21 (0.903)
0.125	24:	03.16 (0.743)	11.03 (0.777)	00.11 (0.841)	01.60 (0.759)	07.54 (0.834)	00.11 (0.841)
0.968	25:	07.04 (0.744)	05.96 (0.830)	06.83 (0.773)	38.90 (0.879)	14.55 (0.868)	38.90 (0.879)
0.410	26:	08.03 (0.828)	05.45 (0.822)	16.47 (0.747)	19.90 (0.728)	01.60 (0.798)	08.03 (0.828)
0.116	27:	17.33 (0.824)	26.50 (0.800)	08.15 (0.770)	18.06 (0.790)	04.52 (0.871)	04.52 (0.871)
0.970	28:	15.67 (0.760)	04.01 (0.865)	13.85 (0.857)	06.56 (0.876)	13.96 (0.847)	06.56 (0.876)
0.586	29:	27.59 (0.798)	02.62 (0.869)	12.55 (0.756)	22.49 (0.782)	02.45 (0.820)	02.62 (0.869)
0.822	30:	01.89 (0.801)	08.85 (0.873)	15.88 (0.777)	01.92 (0.853)	02.17 (0.863)	08.85 (0.873)
0.711	31:	00.34 (0.762)	00.95 (0.831)	03.53 (0.746)	00.96 (0.775)	04.28 (0.833)	04.28 (0.833)
0.769	32:	01.87 (0.817)	00.49 (0.777)	02.44 (0.778)	00.89 (0.883)	26.37 (0.891)	26.37 (0.891)
0.740	33:	01.78 (0.825)	00.39 (0.823)	02.21 (0.784)	02.37 (0.858)	01.55 (0.803)	02.37 (0.858)
0.755	34:	00.33 (0.769)	01.38 (0.793)	00.21 (0.792)	00.92 (0.865)	01.04 (0.818)	00.92 (0.865)
0.748	35:	00.17 (0.724)	01.00 (0.800)	01.28 (0.843)	01.51 (0.774)	04.99 (0.877)	04.99 (0.877)
0.751	36:	31.87 (0.827)	00.70 (0.776)	07.58 (0.835)	00.92 (0.854)	03.08 (0.802)	00.92 (0.854)
0.251	37:	25.32 (0.791)	19.73 (0.744)	08.63 (0.854)	22.32 (0.805)	16.99 (0.807)	08.63 (0.854)
0.933	38:	22.68 (0.812)	12.59 (0.810)	00.14 (0.799)	00.37 (0.858)	01.31 (0.793)	00.37 (0.858)
0.630	39:	00.30 (0.757)	19.78 (0.775)	12.45 (0.816)	04.25 (0.862)	08.02 (0.828)	04.25 (0.862)

Ekspertki wybór najlepszych ?



Rząd:		4	5	6	7	8	Najlepsze
0.223	20	:0.75	2.80	0.30	1.45	0.75	0.75
0.059	21	:0.75	1.00	0.15	0.15	5.10	0.15
0.985	22	:1.45	0.60	0.35	0.40	0.45	0.35
0.439	23	:0.10	0.45	0.40	0.25	1.85	0.10
0.125	24	:0.35	0.00	0.11	0.65	1.05	0.00
0.968	25	:0.50	3.35	0.25	0.35	0.90	0.50
0.410	26	:0.05	0.10	1.10	0.75	1.60	0.05
0.116	27	:2.95	0.70	2.00	3.90	0.85	0.70
0.970	28	:1.45	0.10	0.60	2.85	0.25	0.10
0.586	29	:0.05	0.50	0.05	0.45	1.10	0.05
0.822	30	:0.10	0.75	0.05	0.55	0.35	0.10
0.711	31	:0.15	0.30	0.10	0.15	0.50	0.10
0.769	32	:0.05	0.40	0.20	0.00	0.00	0.00
0.740	33	:0.05	0.05	0.05	0.10	0.00	0.00
0.755	34	:0.00	0.00	0.20	0.15	0.10	0.00
0.748	35	:0.10	0.00	0.15	0.25	0.05	0.00
0.751	36	:0.00	0.10	0.10	0.05	0.40	0.00
0.251	37	:1.80	3.60	1.25	5.85	0.50	0.50
0.933	38	:2.50	0.10	0.14	0.35	1.30	0.10
0.630	39	:0.25	2.00	0.40	0.20	1.10	0.20

Skąd rozbieżności ?



Rząd wielomianu - 5 (z innej serii testów)

0.751 16 :

```
best:26:p=0.750 (85.8%|0.766afit|0c),c 1 at 16 node 6 TP=4 rkg=0.659786 tCheat=3
wins:30:p=0.468 (98.1%|0.858afit|0c),c 1 at 16 node 5 TP=1 rkg=0.839440 tCheat=0*
wins:30:p=0.468 (98.1%|0.858afit|0c),c 1 at 16 node 5 TP=1 rkg=0.839440 tCheat=0*
wins:28:p=0.468 (98.1%|0.858afit|0c),c 1 at 16 node 5 TP=1 rkg=0.839440 tCheat=0*
wins:28:p=0.468 (98.1%|0.858afit|0c),c 1 at 16 node 5 TP=1 rkg=0.839440 tCheat=0*
wins:27:p=0.468 (98.1%|0.858afit|0c),c 1 at 16 node 5 TP=1 rkg=0.839440 tCheat=0*
wins:27:p=0.468 (98.1%|0.858afit|0c),c 1 at 16 node 5 TP=1 rkg=0.839440 tCheat=0*
wins:27:p=0.468 (98.1%|0.858afit|0c),c 1 at 16 node 5 TP=1 rkg=0.839440 tCheat=0*
wins:28:p=0.873 (85.1%|0.831afit|0c),c 2 at 16 node 3 TP=2 rkg=0.694223 tCheat=2*
wins:32:p=0.873 (85.1%|0.831afit|0c),c 1 at 16 node 3 TP=1 rkg=0.679223 tCheat=1*
```

0.251 17 :

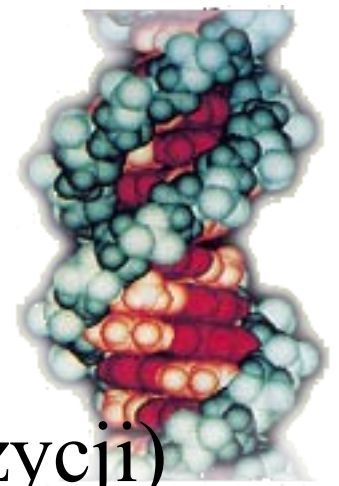
```
best:23:p=0.202 (79.0%|0.767afit|1c),c 1 at 17 node 3 TP=4 rkg=0.057954 tCheat=5
wins:31:p=0.806 (96.0%|0.840afit|0c),c 2 at 17 node 6 TP=2 rkg=0.814919 tCheat=0*
wins:32:p=0.806 (96.0%|0.840afit|0c),c 1 at 17 node 6 TP=1 rkg=0.799919 tCheat=0*
wins:31:p=0.806 (96.0%|0.840afit|0c),c 1 at 17 node 6 TP=1 rkg=0.799919 tCheat=0*
wins:29:p=0.806 (96.0%|0.840afit|0c),c 1 at 17 node 6 TP=1 rkg=0.799919 tCheat=0*
wins:28:p=0.806 (96.0%|0.840afit|0c),c 1 at 17 node 6 TP=1 rkg=0.799919 tCheat=0*
wins:28:p=0.806 (96.0%|0.840afit|0c),c 1 at 17 node 6 TP=1 rkg=0.799919 tCheat=0*
wins:27:p=0.806 (96.0%|0.840afit|0c),c 1 at 17 node 6 TP=1 rkg=0.799919 tCheat=0*
wins:27:p=0.806 (96.0%|0.840afit|0c),c 1 at 17 node 6 TP=1 rkg=0.799919 tCheat=0*
wins:27:p=0.806 (96.0%|0.840afit|0c),c 1 at 17 node 6 TP=1 rkg=0.799919 tCheat=0*
```

Wyniki testów



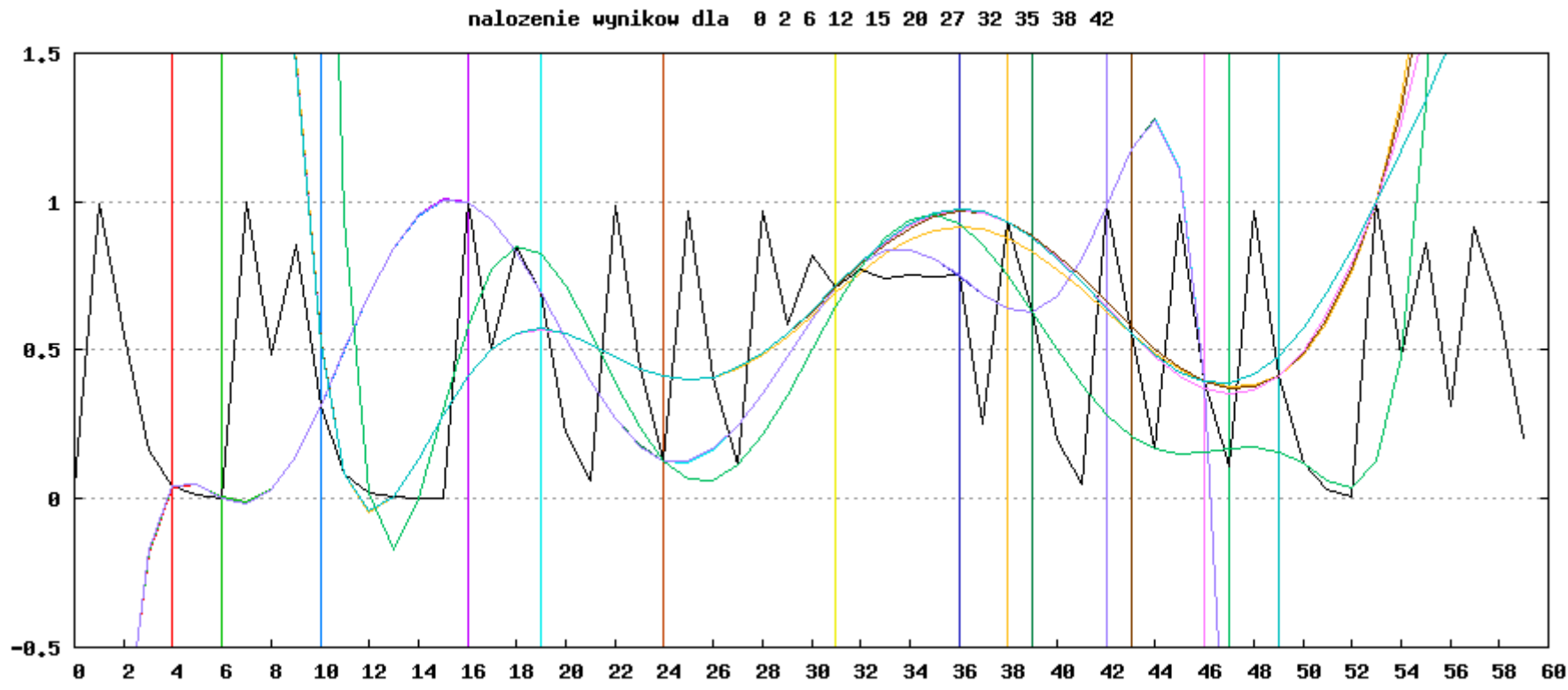
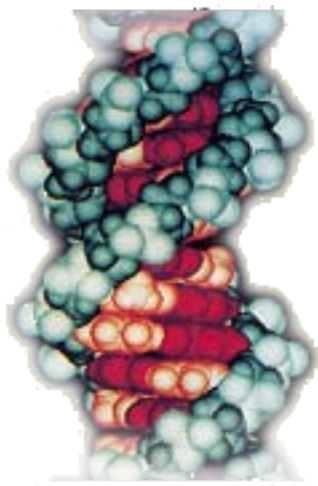
- Wielomiany to silne narzędzie interpolacyjne, być może cały proces wyboru węzłów nic lub niewiele wnosi do interpolacji
- Warto sprawdzić jakie wyniki daje interpolacja z użyciem węzłów symetrycznie rozłożonych względem danej brakującej
- Interpolacja została sprawdzona na przedziale $[0,39]$ testowanego szeregu logistycznego

Uczenie analiza

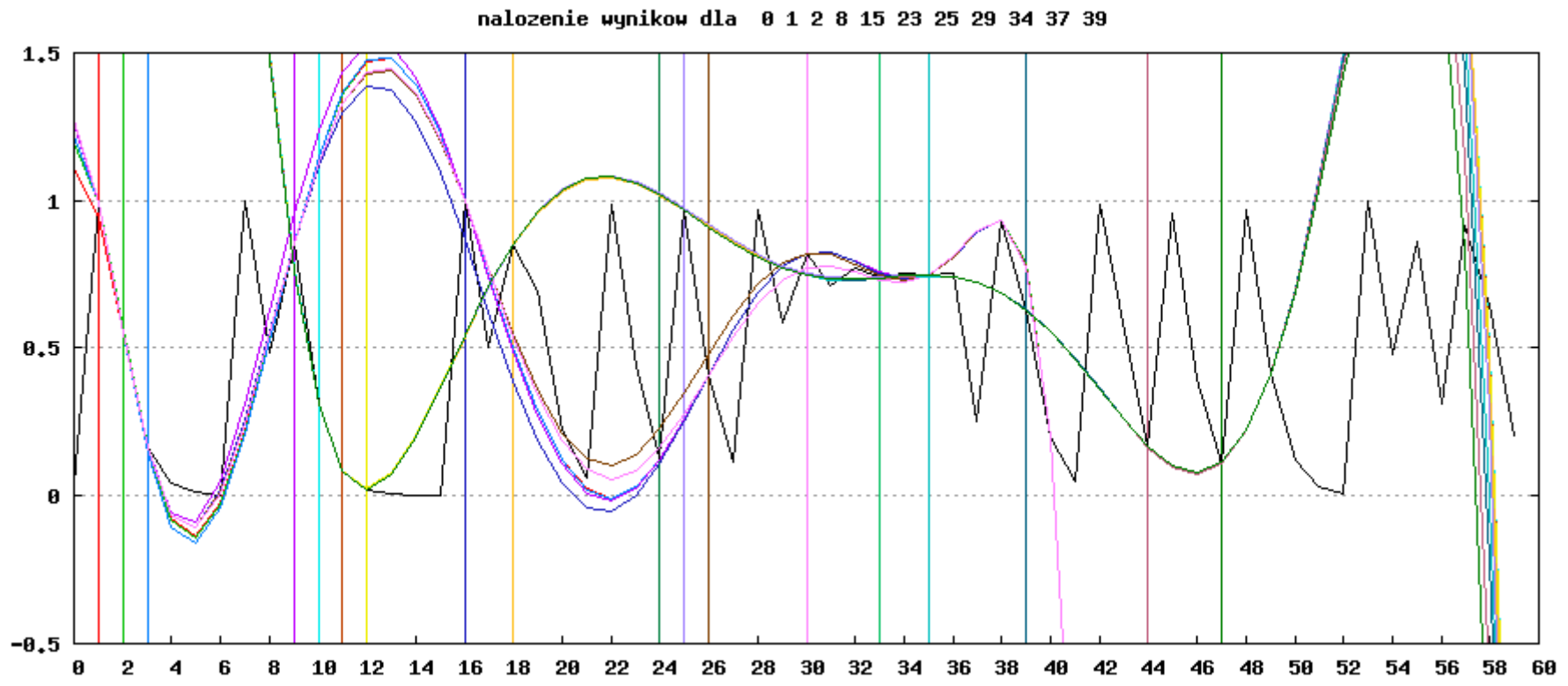
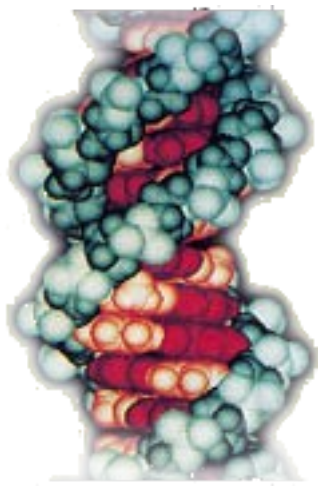


- Uczenie na pełnym zbiorze danych (60 pozycji)
- Jak wyglądają wielomiany uczące (dopasowane do danych uczących), czy są podobne kształtem ?
- Ile pozycji pokrywa wielomian ?
- Jaki procent danych wejściowych pokrywa n najlepszych wielomianów z populacji ?
- Nałożenie najważniejszych wielomianów

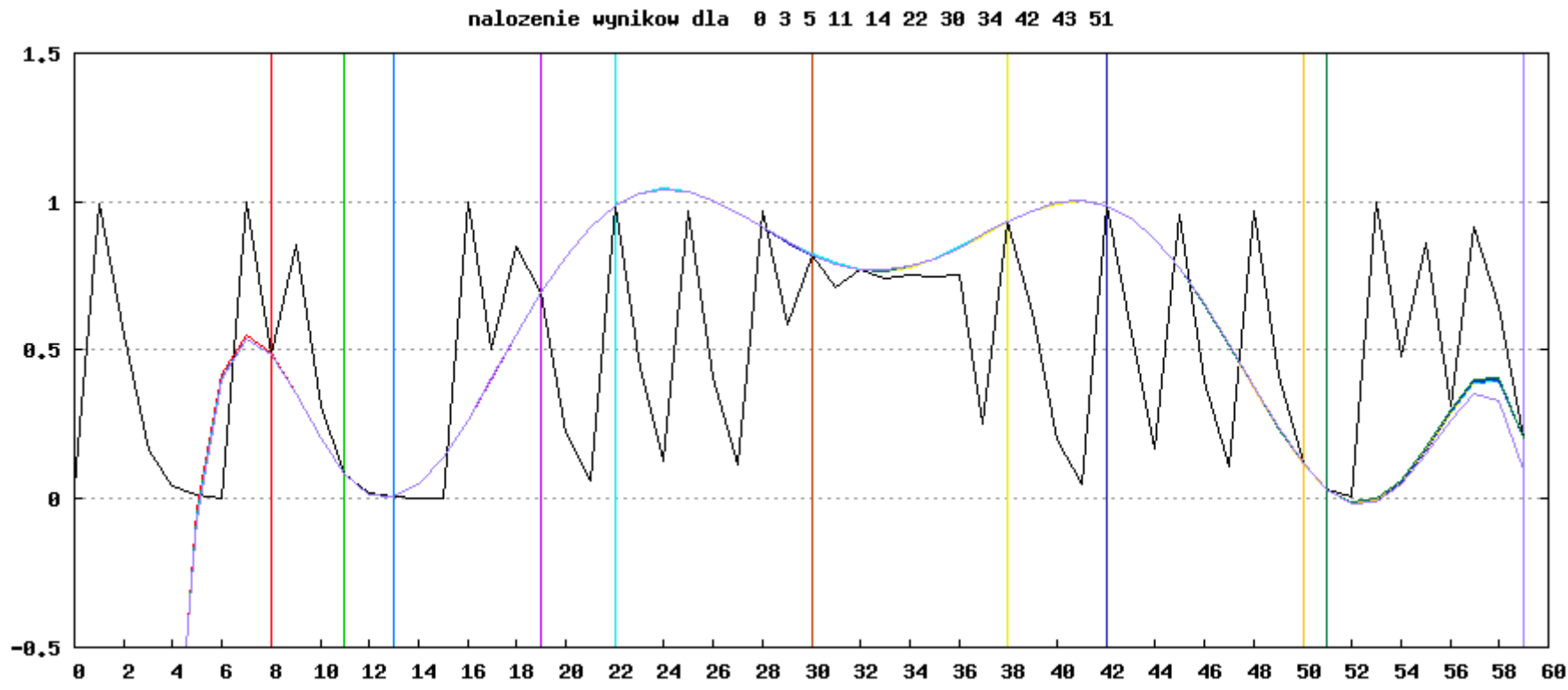
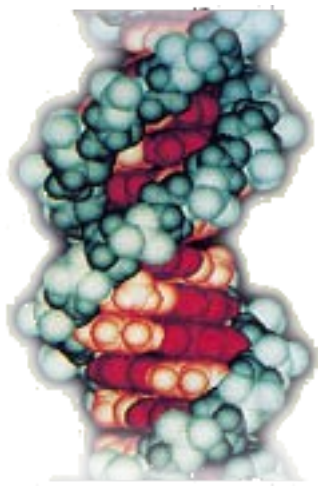
Punkty budowania oceny osobnika



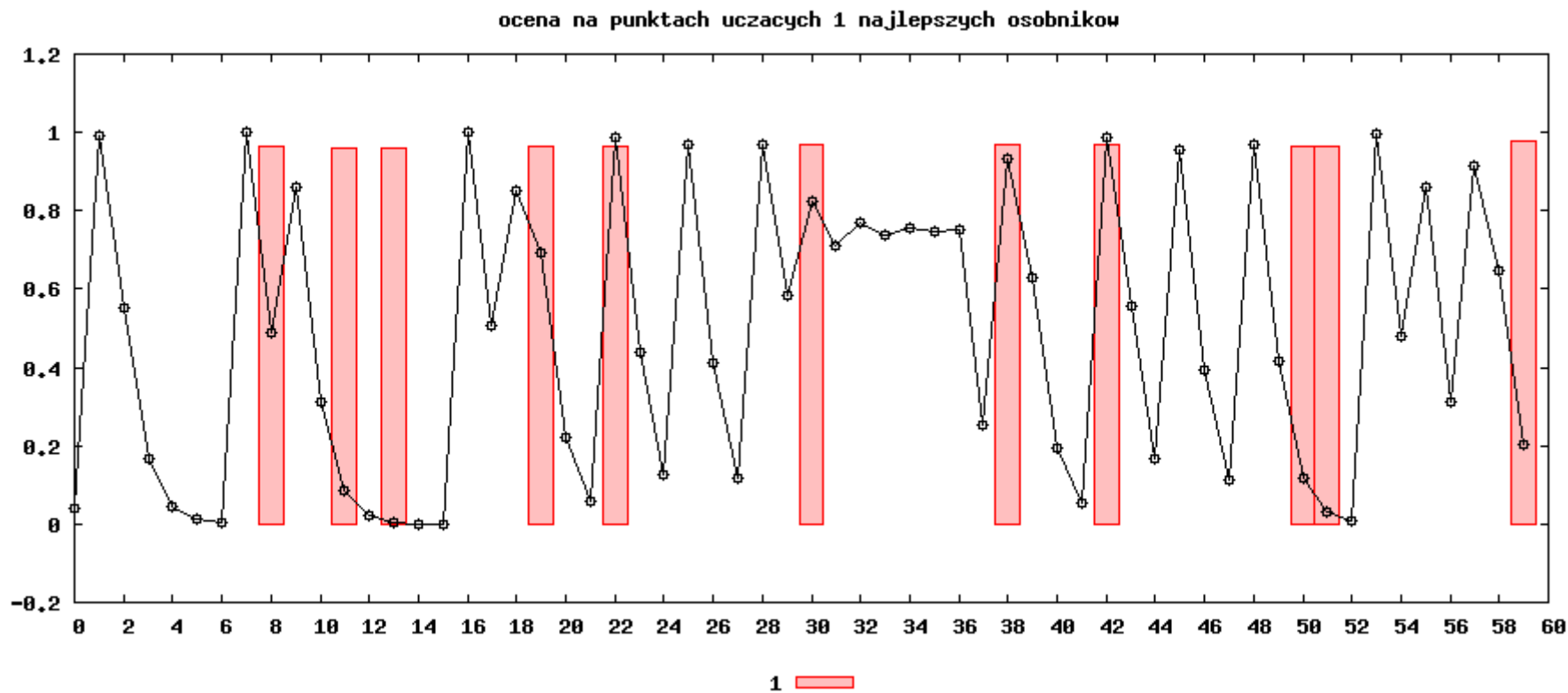
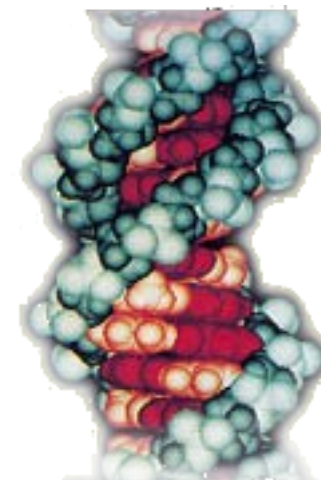
Punkty budowania oceny osobnika



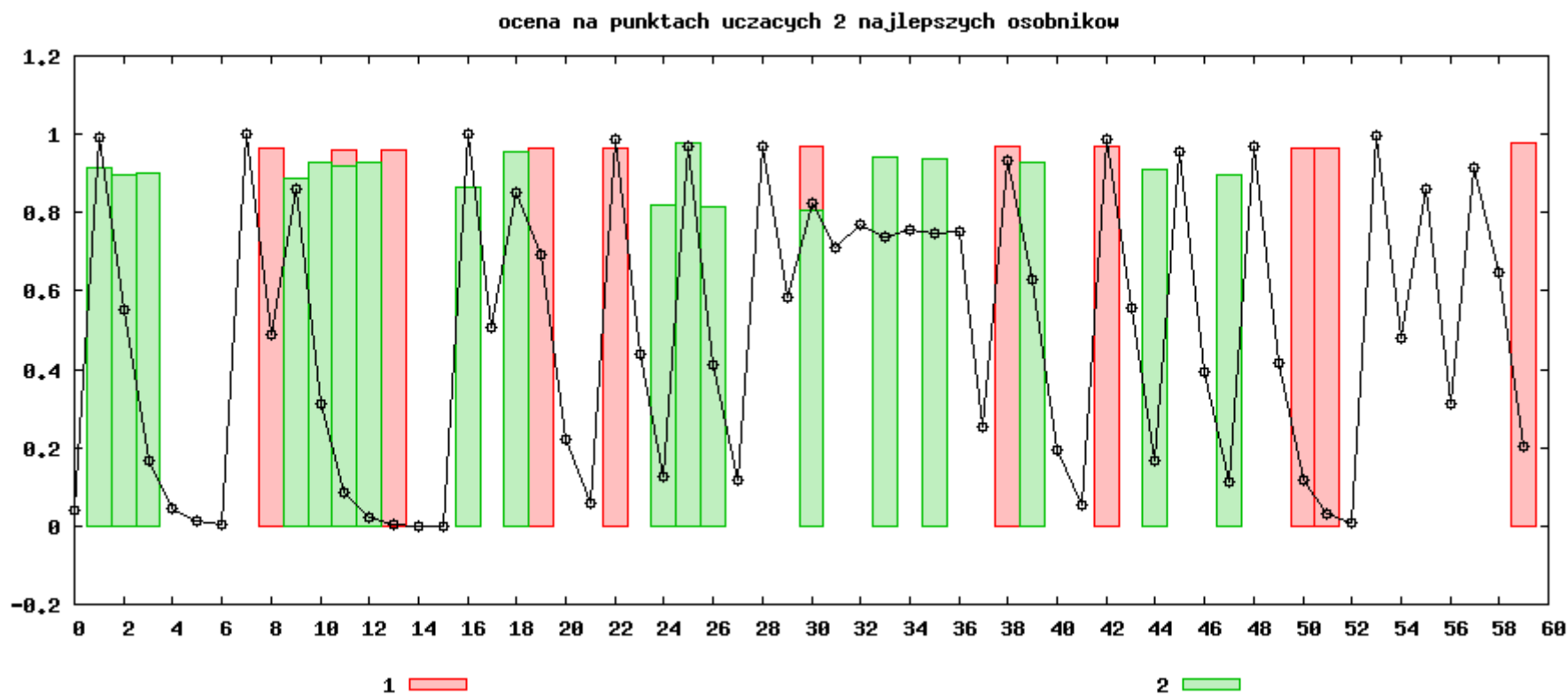
Punkty budowania oceny osobnika



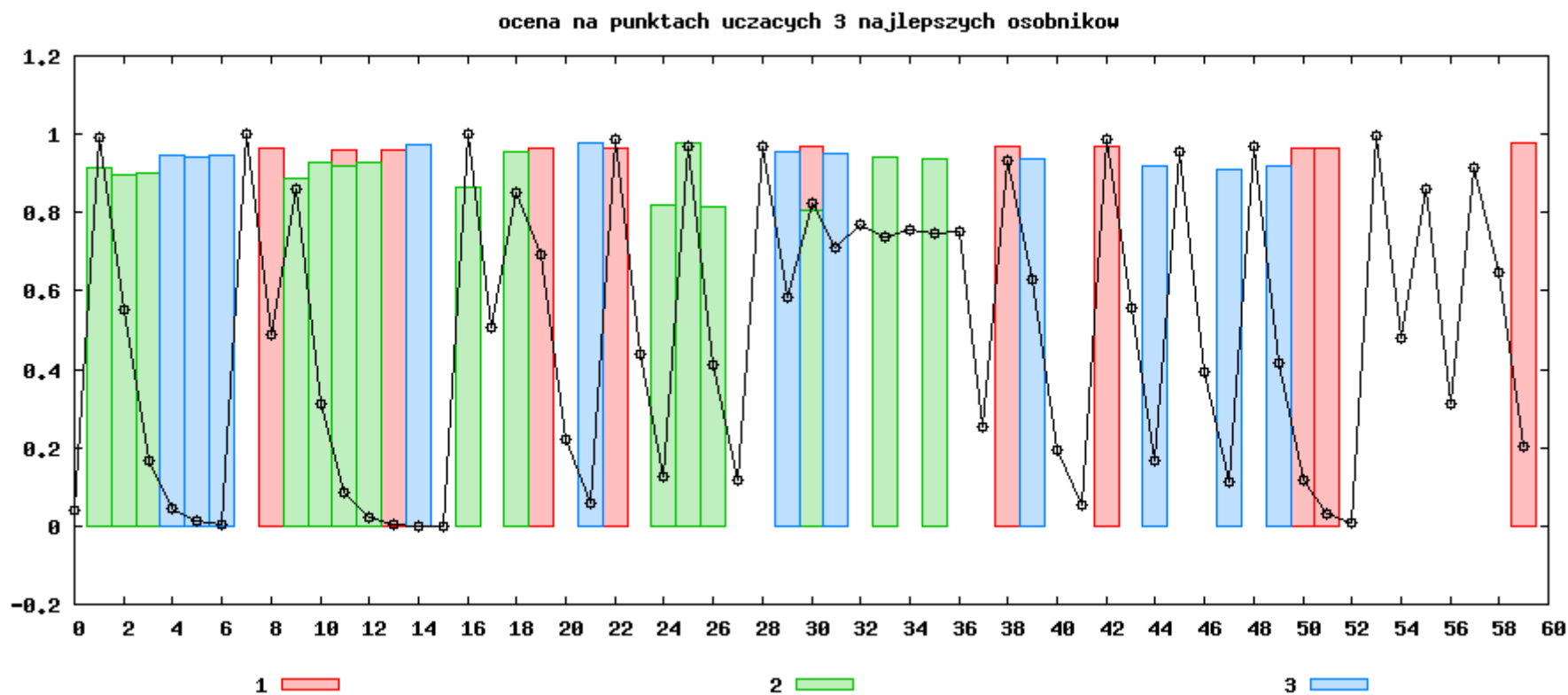
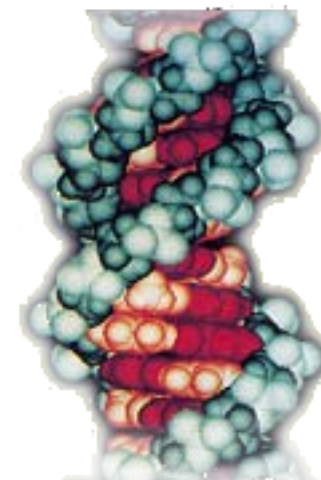
Pokrycie obszaru wielomianami



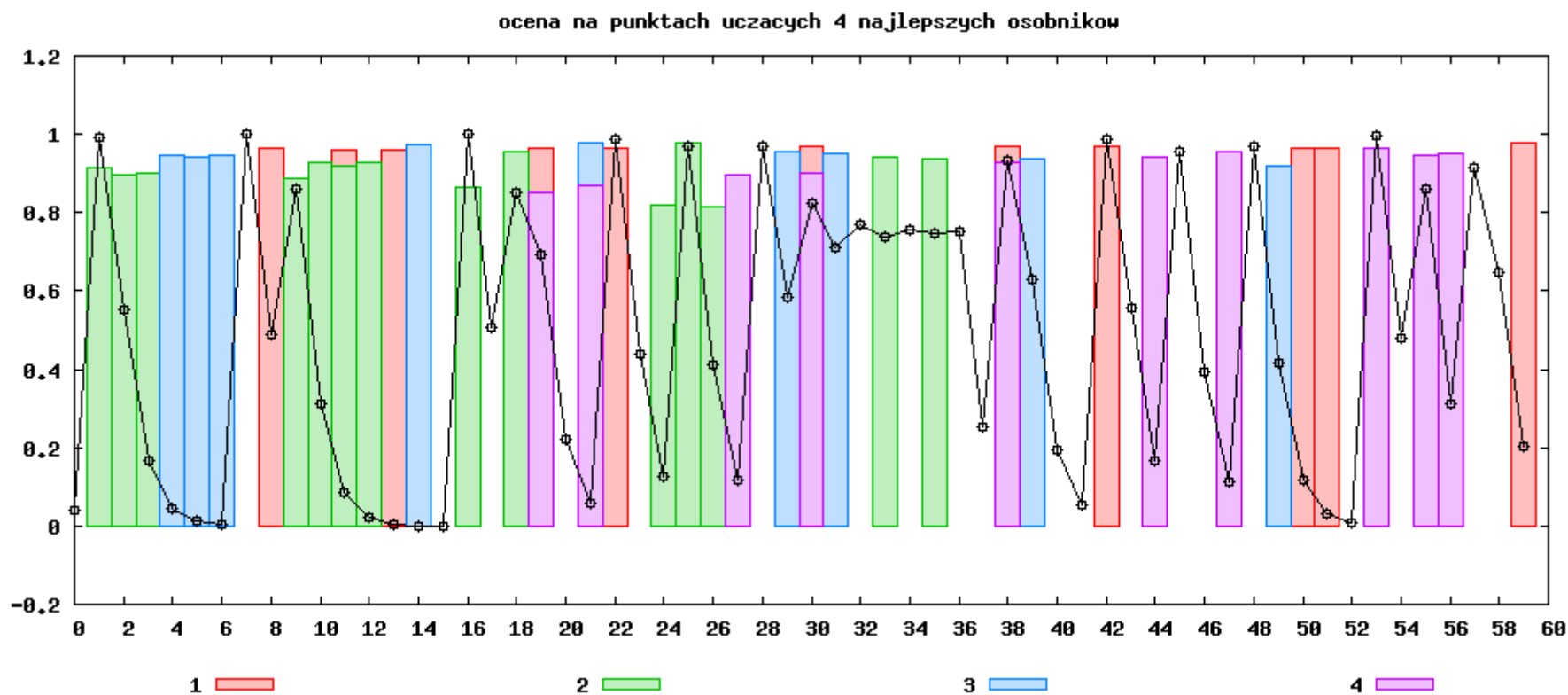
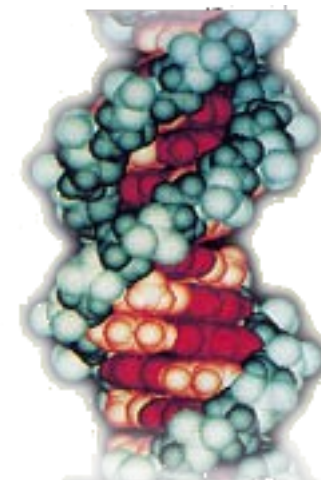
Pokrycie obszaru wielomianami



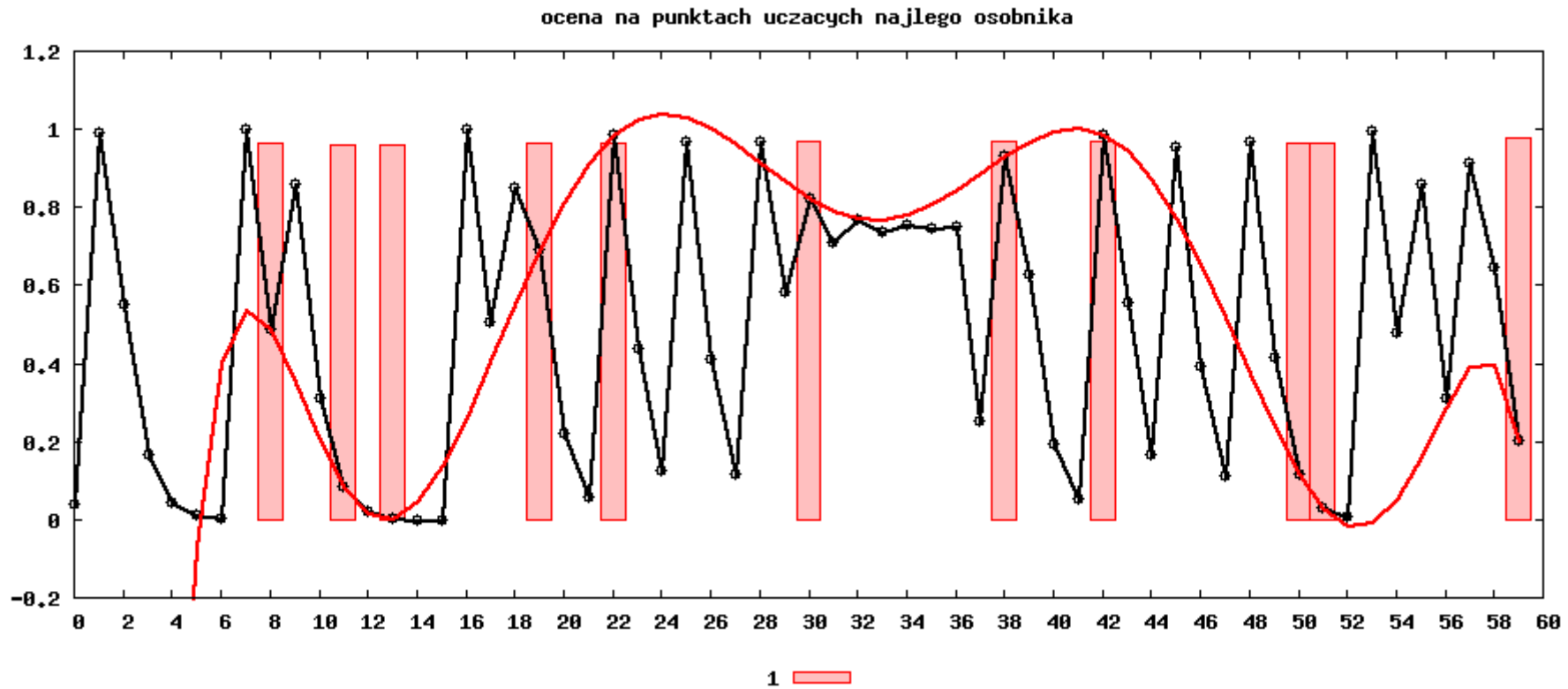
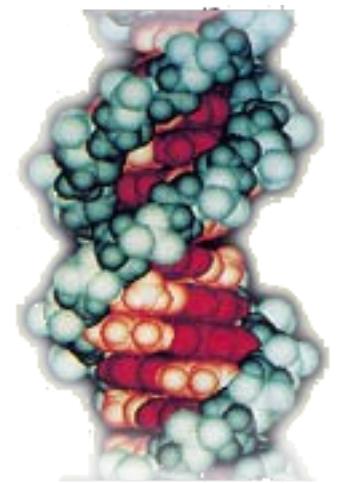
Pokrycie obszaru wielomianami



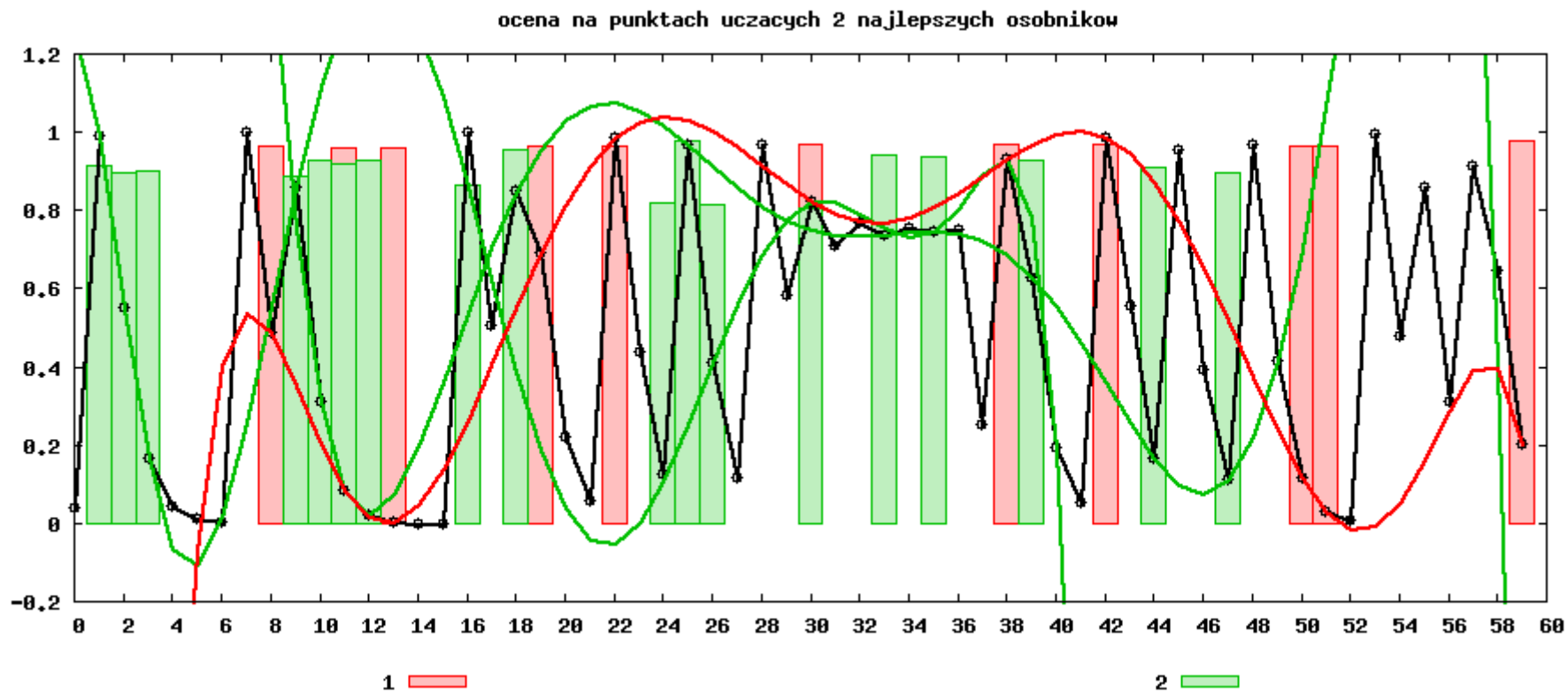
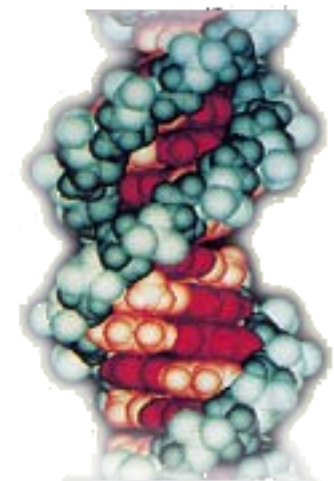
Pokrycie obszaru wielomianami



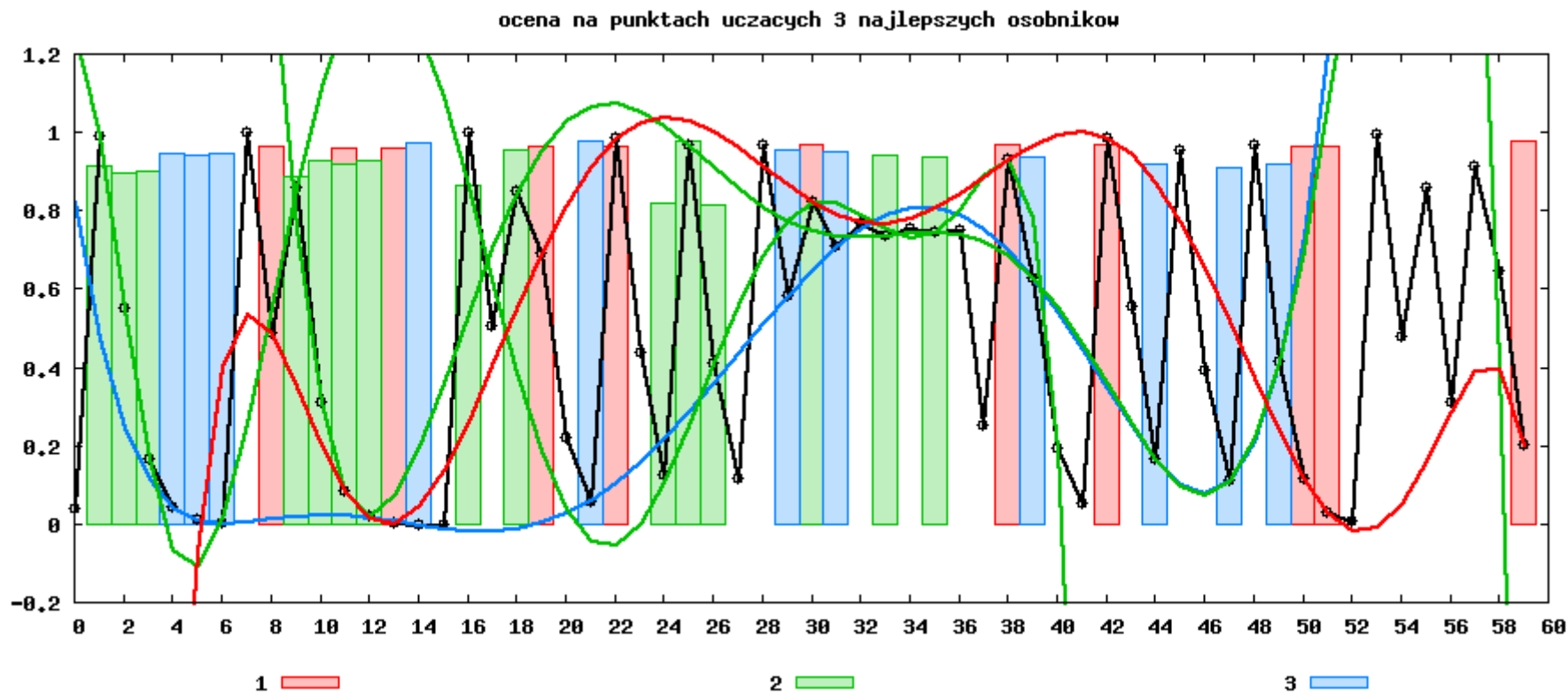
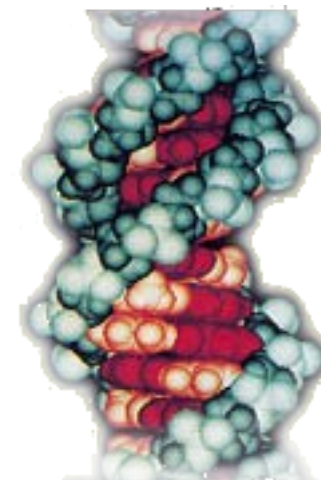
Nałożenie najlepszych osobników



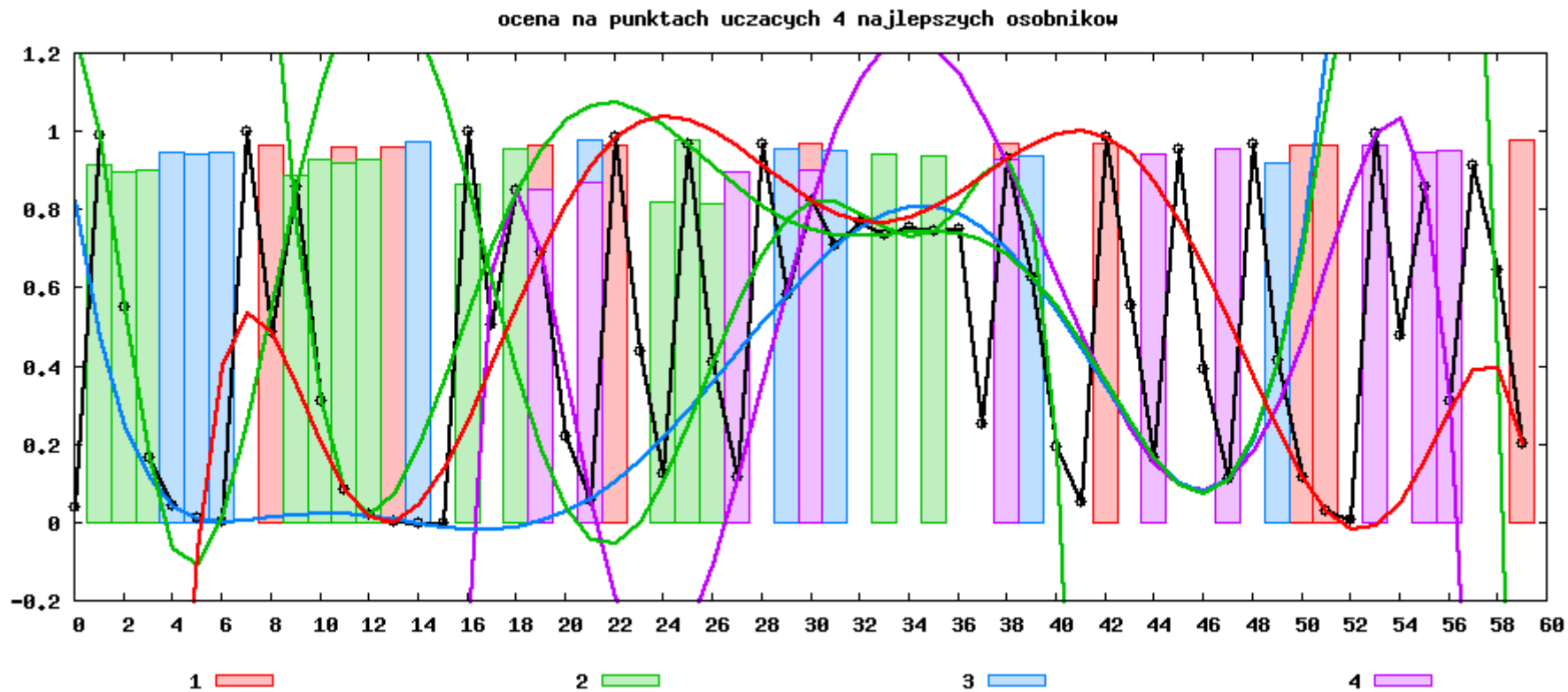
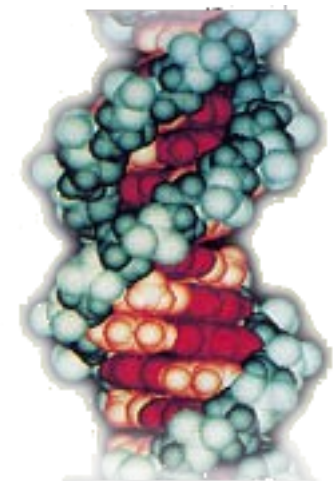
Nałożenie najlepszych osobników



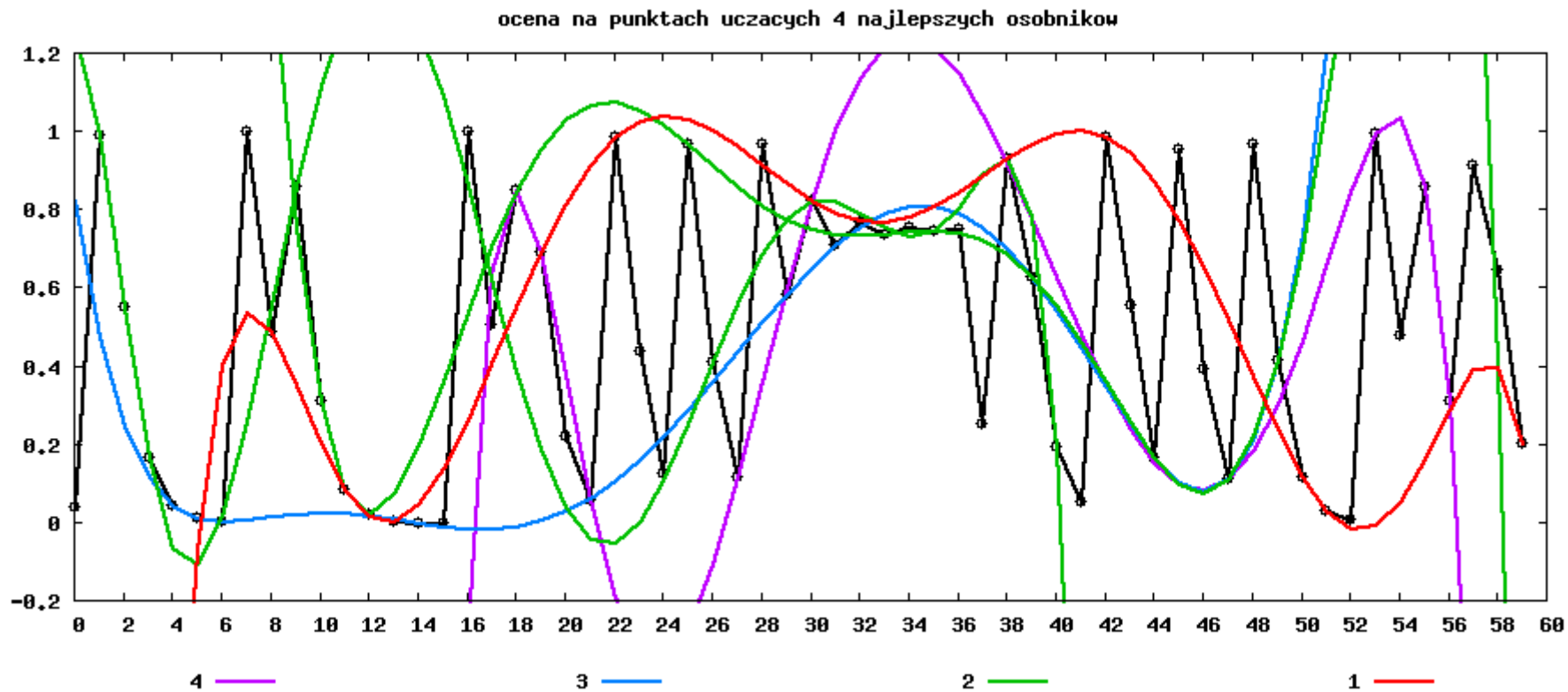
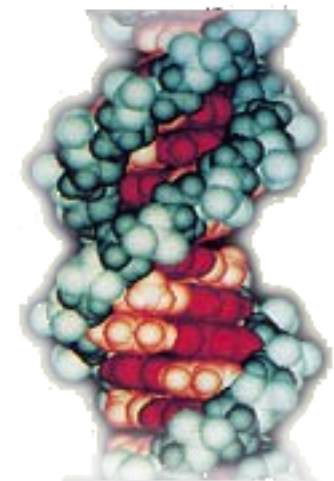
Nałożenie najlepszych osobników



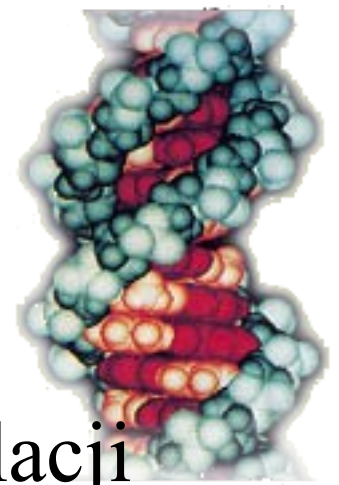
Nałożenie najlepszych osobników



Nałożenie najlepszych osobników

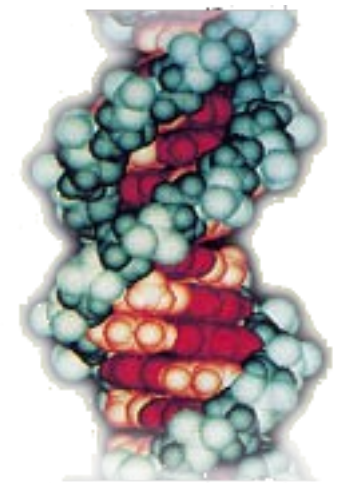


Dalsze plany



- Wielomiany różnych stopni w jednej populacji
- Optymalizacja czasu testów
- Jeszcze jeden dodatkowy węzeł do weryfikacji
- “Usztywnienie” kształtu wielomianu
- Testy na większym odcinku danych
- Uczenie na jednym odcinku, uzupełnianie na innym
- Zostawić interpolacje – wyszukiwanie reguł w danych

To już wszystko!



Dziękuję za uwagę,
proszę o pytania i komentarze.

Marcin Borkowski
marcinbo@mini.pw.edu.pl