

Optymalizacja metodą algorytmów genetycznych - lab03

Tomasz Rajchel (297908)

Informatyka Stosowlana WFILS

2018/10/26

Cel laboratorium

Znalezienie maksimum funkcji $f(x) = x \sin(10\pi * x) + 1$ w przedziale $x \in [-1, 2]$, reprezentowane przez ciąg 22 bitów. Następnie analiza różnych metod kodowania osobników.

Algorytm genetyczny

Genotyp każdego osobnika naszej populacji to ciąg 22 bitów które na początku są generowane losowo. Koduje on chechy naszego osobnika. Jest on zakodowany w naturalnym kodzie binarnym $(0 - 4194304)_{10}$, a potem rzutowany na liczby rzeczywiste w przedziale $[-1, 2]$, czyli dziedzinę naszego x .

Testujemy algorytm dla populacji 20 osobników. Funkcja przystosowania dla każdego osobnika jest zdefiniowana następująco:

$$f_1(x) = x \sin(10\pi * x) + 2$$

(+2, a nie +1 żeby była nieujemna w swojej dziedzinie, ułatwia to obliczenia)

Po każdej epoce osobniki są ewaluowane funkcją przystosowania. Następnie dokonujemy selekcji metodą ruletki tzn. prawdopodobieństwo wyboru osobnika do krzyżowania jest proporcjonalne do jego wartości funkcji przystosowania. Dopuszczamy możliwość krzyżowania osobnika z samym sobą, wtedy jest on po prostu klonowany.

Krzyżowanie odbywa się w sposób jednopunktowy. Wybieramy losowo punkt w genotypie (locus) i przepisujemy 2 osobniki „na krzyż”, tworząc 2 nowe osobniki.

Dodatkowo osobniki podlegają mutacji. Losowy bit w genotypie zostaje przestawiony z 1 na 0 lub odwrotnie. Mutacja zachodzi dla każdego genu z tym samym prawdopodobieństwem $p_{mutation}$.

Działanie algorytmu

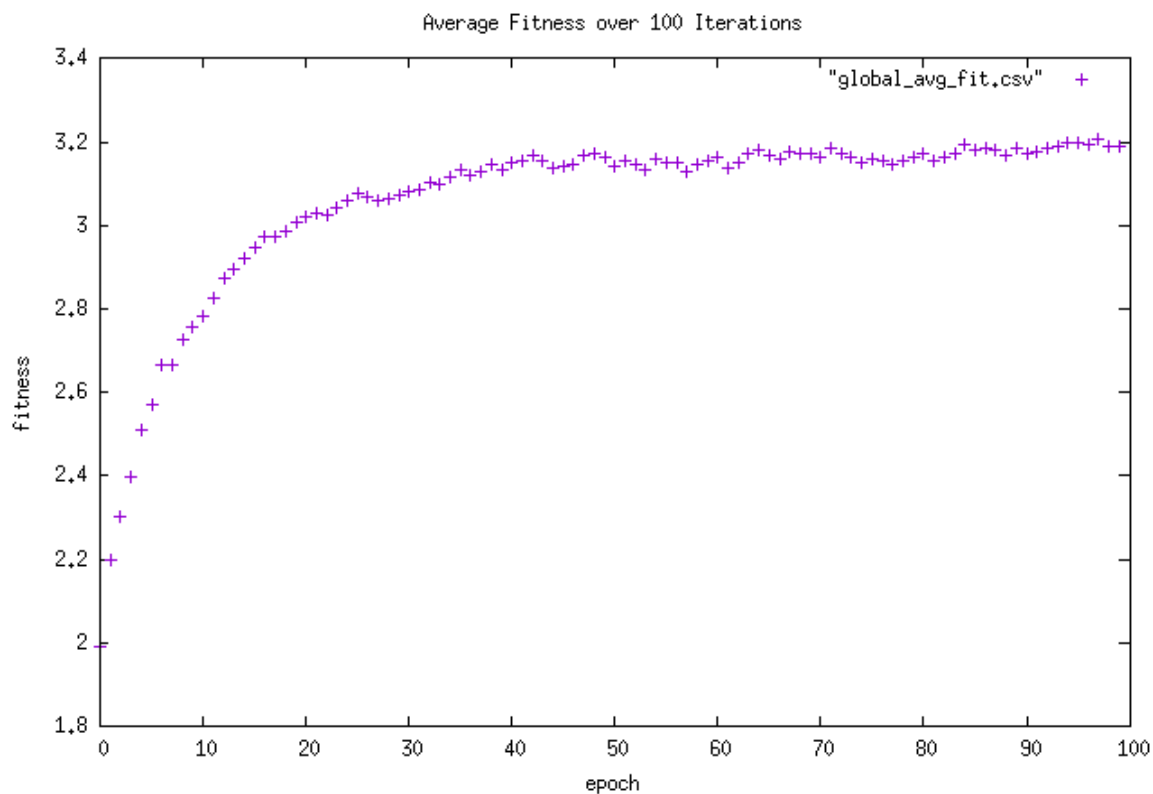
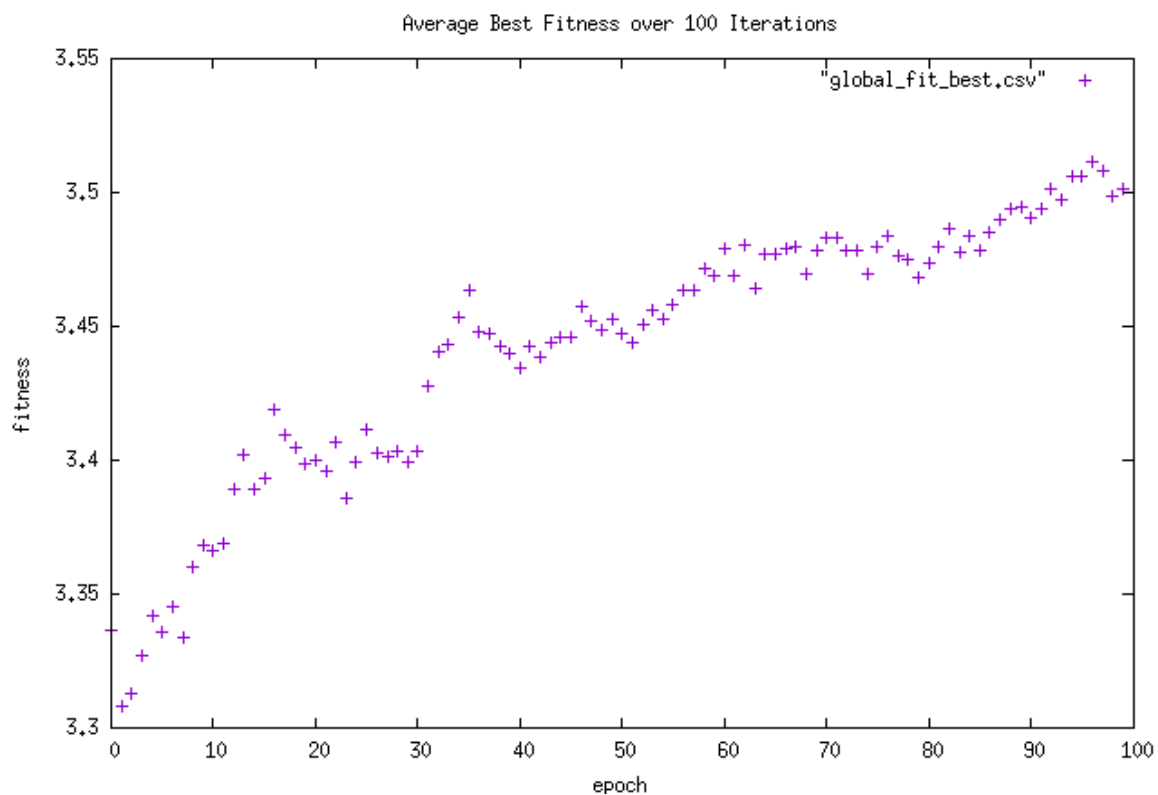
Zobaczmy jak działa algorytm dla następujących parametrów:

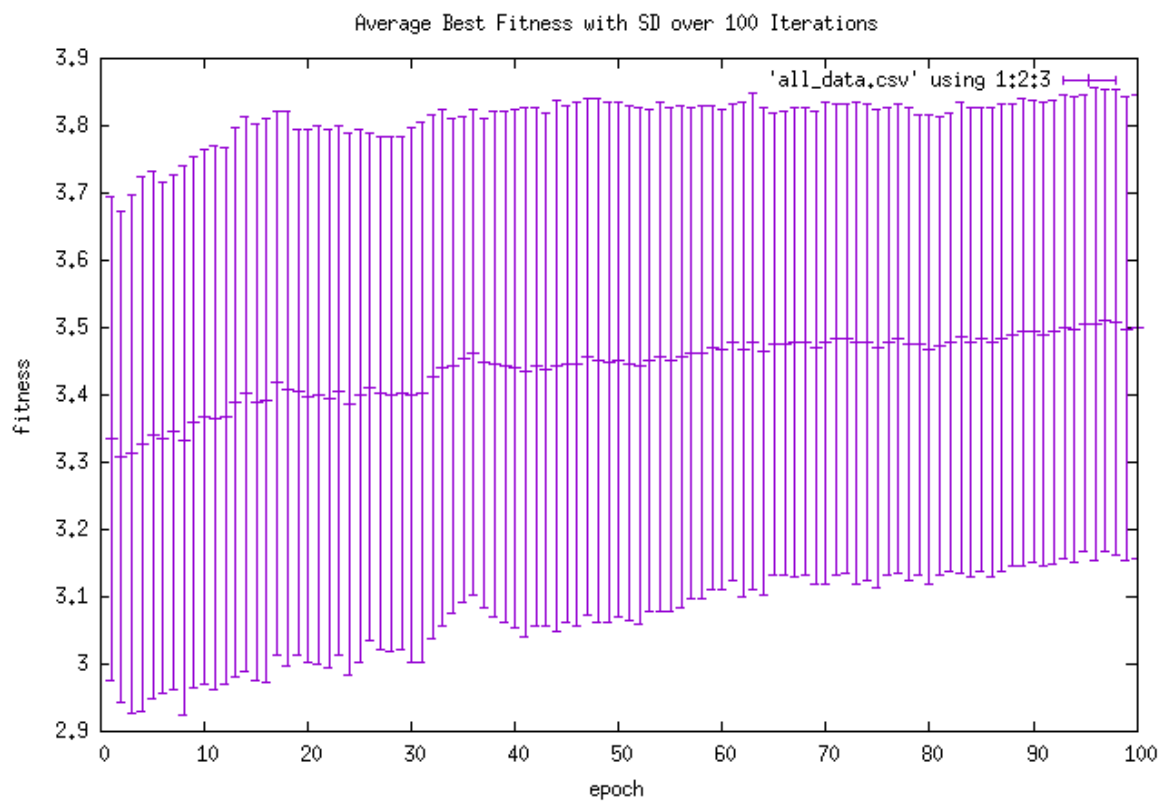
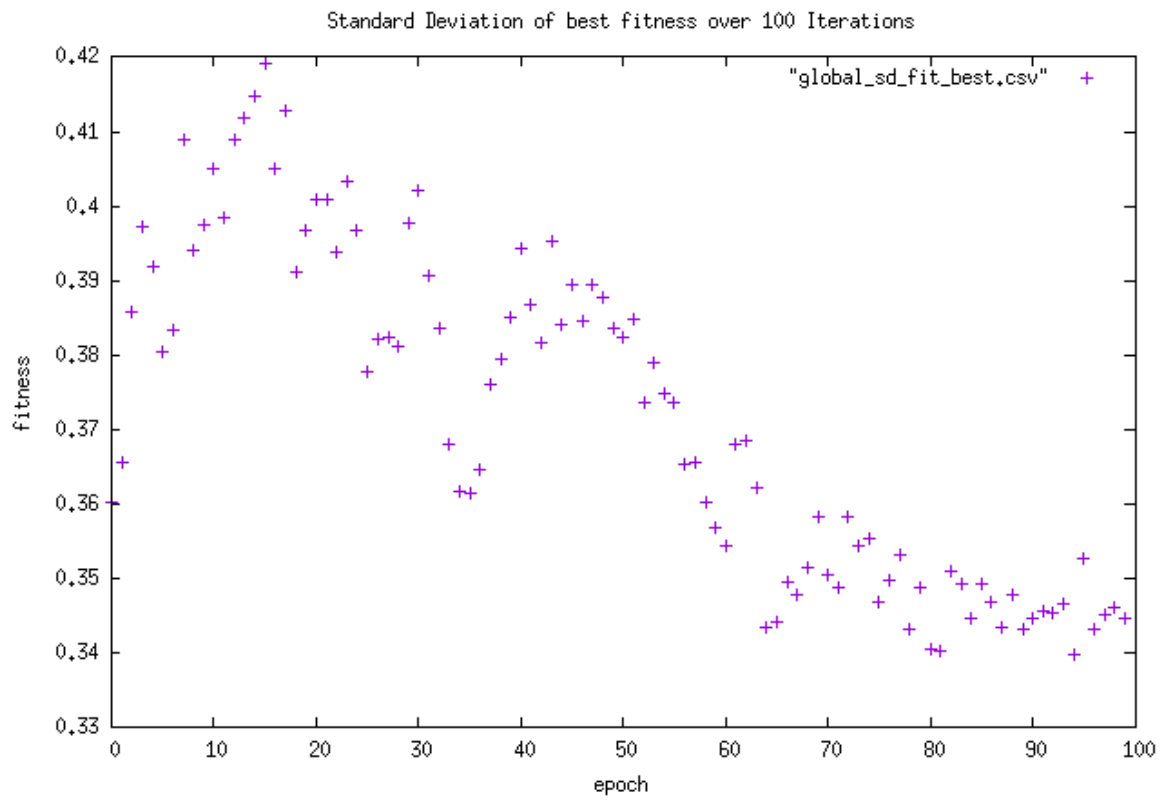
populacja = 20 osobników

$p_{cross} = 1$ (rozmiar populacji nie ulega zmianie)

$p_{mutation} = 0.01$

ilość epok = 100





Należy pamiętać, że algorytm jest niedeterministyczny więc porównujemy średnie wartości ze 100 uruchomień o długości 100 epok.

Najlepsze rozwiązanie ze 100 uruchomień algorytmu to:

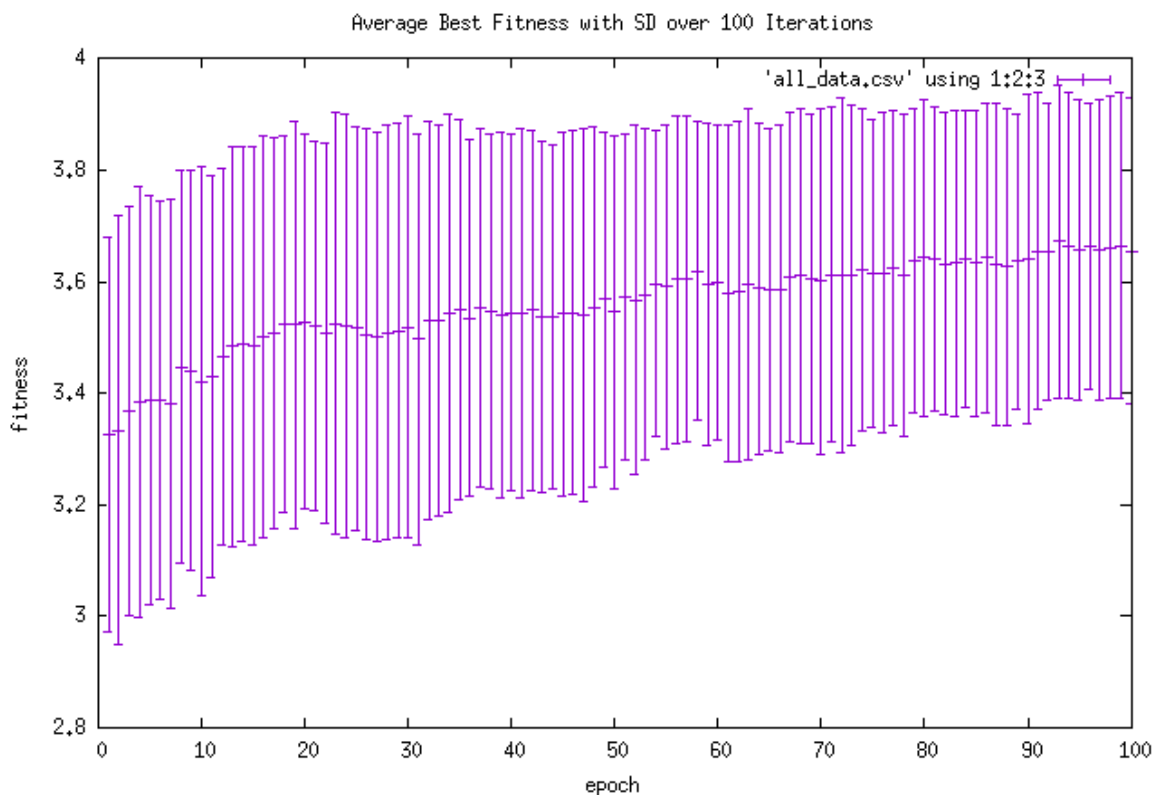
```
best fitness = 3.85027 for x = 1.85055
```

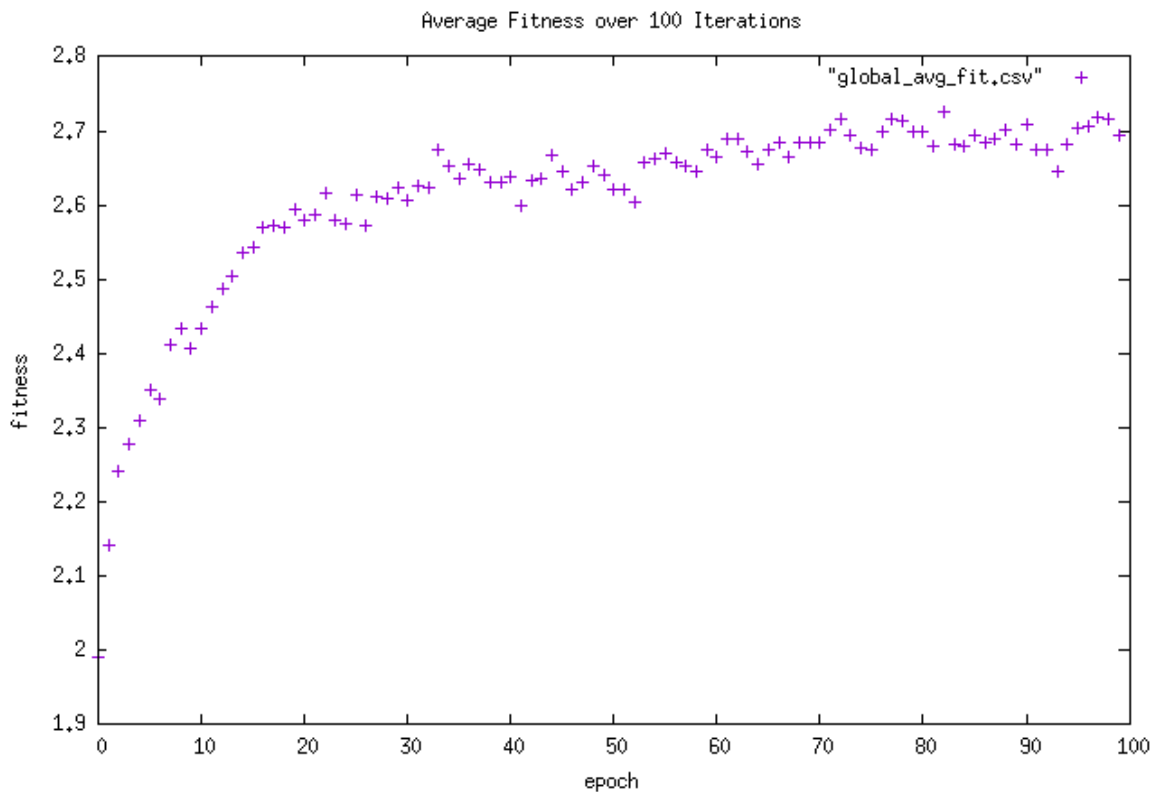
$x = 1.85055$ które dla funkcji $f(x)$ wynosi 2.85027 (3.8 to wartość robocza)

Wraz z kolejnymi epokami :

- Średnia wartość przystosowania najlepszego rozwiązania w populacji rośnie,
- Średnie przystosowanie każdego osobnika także rośnie,
- Odchylenie standardowe przystosowania najlepszego osobnika maleje.

Zobaczmy jak działa algorytm dla $p_{\text{mutation}} = 0.05$ (5 razy większy)

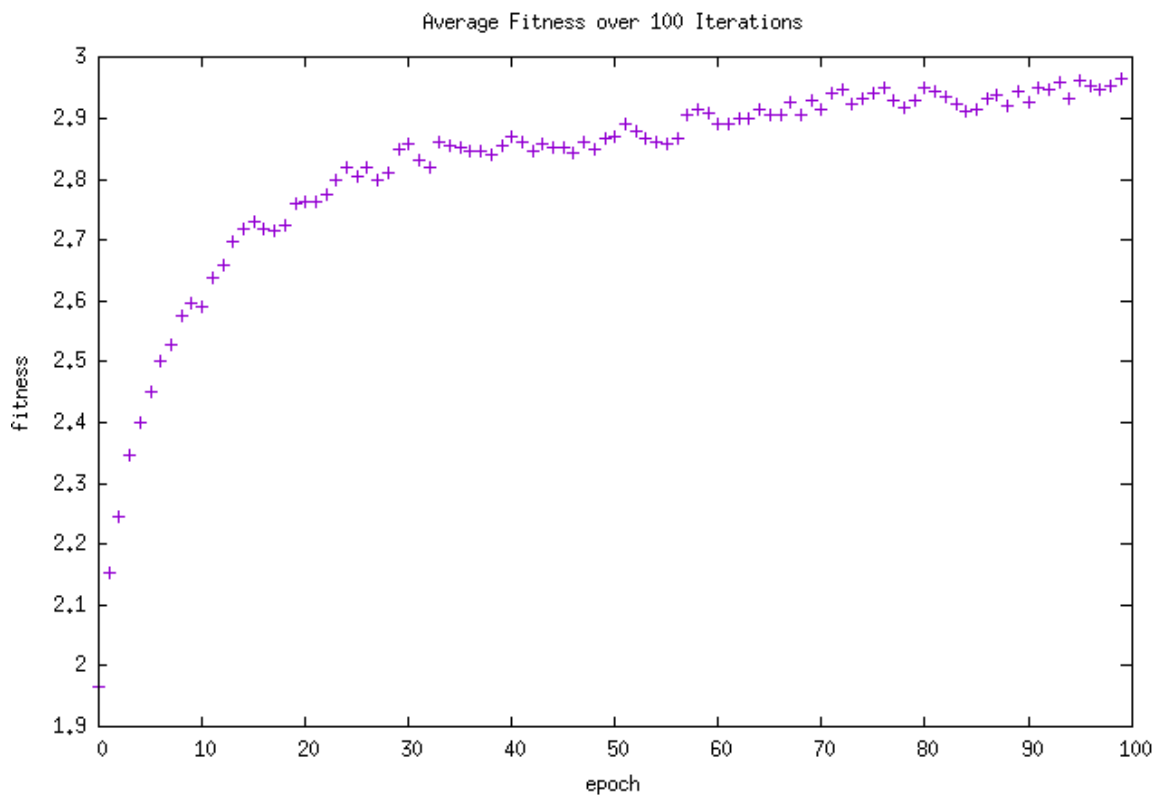
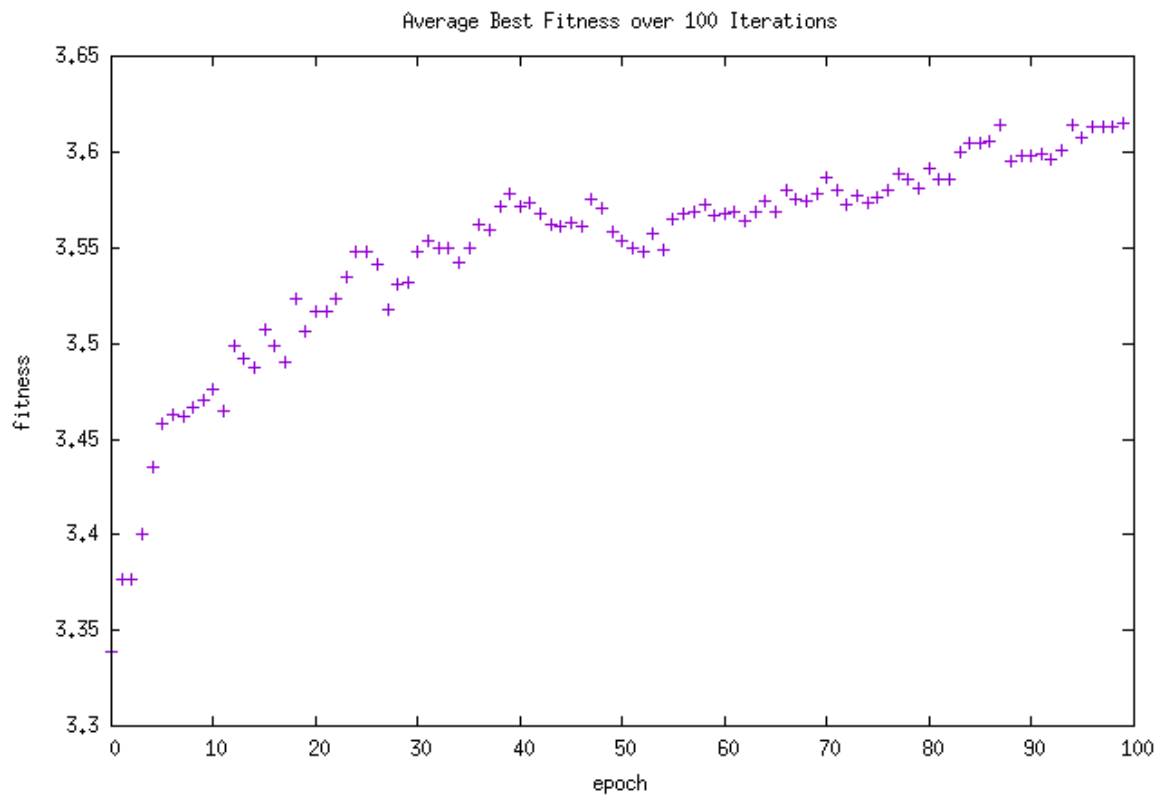


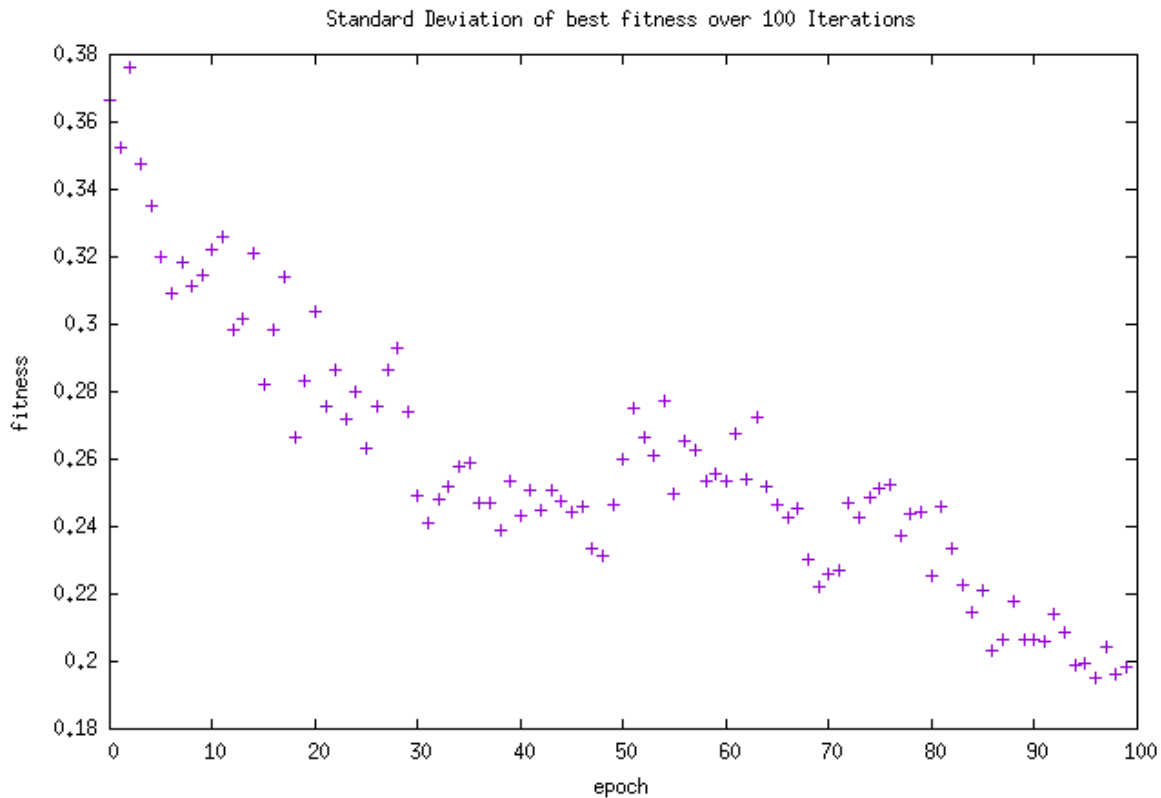


Porównując wyniki z poprzednimi możemy zauważyć, że najlepsze osobniki w populacji są lepsze, ale średnie przystosowanie dla całej populacji jest gorsze. Wśród bardziej zmutowanych osobników łatwiej znaleźć najlepszego, ale cierpi na tym przystosowanie całej populacji.

Kodowanie Graya

Zobaczmy jak zmieni się działanie algorytmu gdy zamiast kodować genotyp naturalnym kodem binarnym zastosujemy kodowanie Graya.



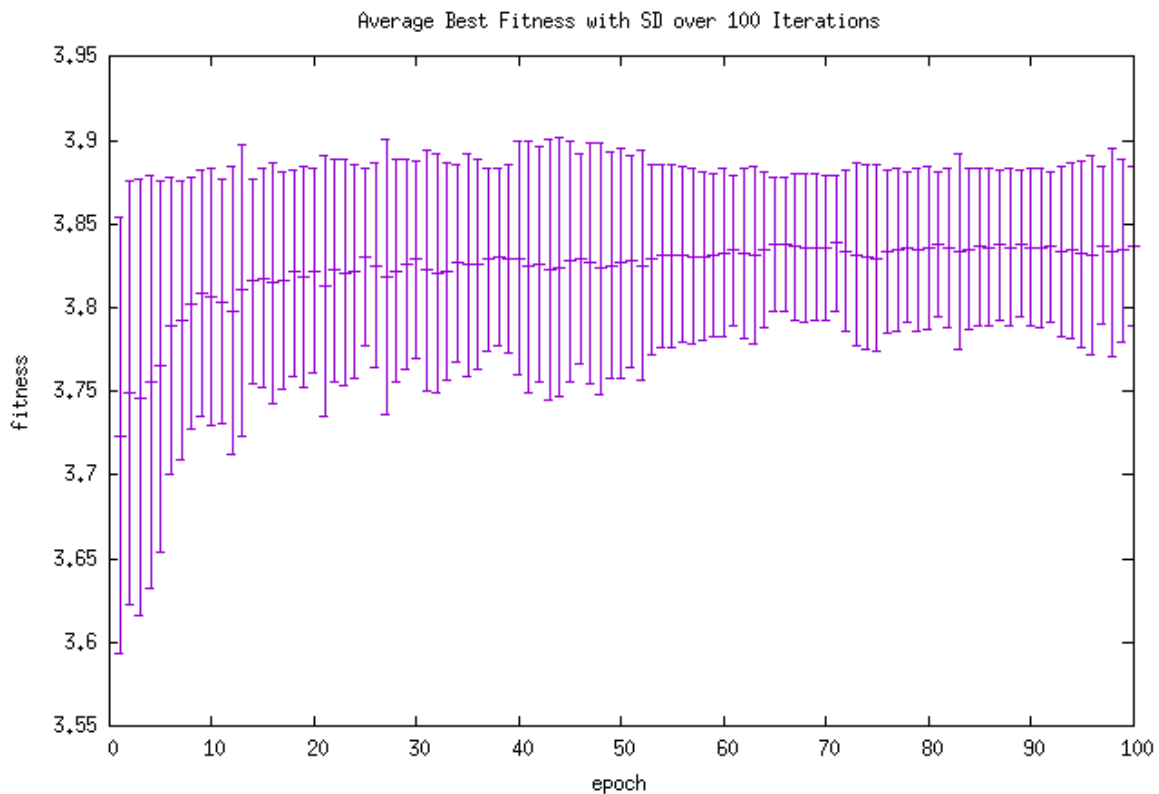


Zauważmy, że stosując kodowanie Graya, odchylenie standardowe (między uruchomieniami) jest o wiele mniejsze (0.19 vs 0.34). Algorytm jest stabilniejszy, nie wpływają już tak bardzo na niego warunki początkowe i mutacje.

Rozmiar populacji

Zwiększenie rozmiaru populacji jeszcze bardziej poprawia stabilność algorytmu (odchylenie standardowe między uruchomieniami jest mniejsze) i poprawia najlepsze rozwiązania, ale zwiększa złożoność obliczeniową, i wymaganą pamięć w sposób liniowy. Nie ma wpływu na średnie przystosowanie.

Poniżej wykres najlepszych rozwiązań, ze standardowym odchyleniem między uruchomieniami dla populacji 80 osobników.



Dalsze ulepszenia

Do populacji w następnej epoce można przepisać najlepszego osobnika bez zmian. Jest to tak zwana strategia elitarna. Istnieją lepsze metody selekcji osobników do krzyżowania niż metoda ruletki, na przykład:

- Selekcja turniejowa
Polega na podzieleniu populacji na grupy i „rozgrywaniu turnieju” między osobnikami wewnątrz grup. Do krzyżowania wybierane są najlepsze osobniki z każdej grupy
- Selekcja rankingowa
Polega na wyborze osobników zgodnie z przypisanymi im rangami. Osobnikom nadawana jest ranga (zależna od wartości przystosowania). Następnie wybiera się osobniki przechodząc w dół listy (dla niższych osobników prawdopodobieństwo jest coraz mniejsze)

Alternatywnie zamiast krzyżowania jednopunktowego można zastosować krzyżowanie wielopunktowe, które jest szybsze dla dłuższych chromosomów albo krzyżowanie równomierne w którym losujemy wzorzec (maskę bitową) według którego przepisujemy genotypy.