



Uczenie maszynowe
- Algorytmy genetyczne

Inteligentne systemy decyzyjne

Optymalizacja

W praktyce inżynierskiej często zachodzi potrzeba znalezienia parametrów, dla których system/urządzenie będzie działać w sposób optymalny. Klasyczne podejście do optymalizacji:

- sformułowanie funkcji celu (zależnej od n zmiennych)
- Poszukiwanie minimum lub maksimum funkcji celu.

Optymalizacja

Problemy związane z optymalizacją w klasycznym ujęciu:

- Skomplikowany model, trudny do opisania matematycznie,
- Potrzeba wykonania dużej liczby obliczeń,
- „Pułapka” minimum lokalnego.

Metody ewolucyjne

Metody ewolucyjne proponują alternatywne podejście do problemu optymalizacji. Dzięki ich zastosowaniu możliwe jest szybsze przeszukiwanie przestrzeni rozwiązań z uniknięciem pułapek minimum lokalnego. Algorytmy ewolucyjne stanowią istotny dział nauki o sztucznej inteligencji. Należy jednak pamiętać, że ich wynik jest jedynie **przybliżeniem** najlepszego rozwiązania.

Definicja

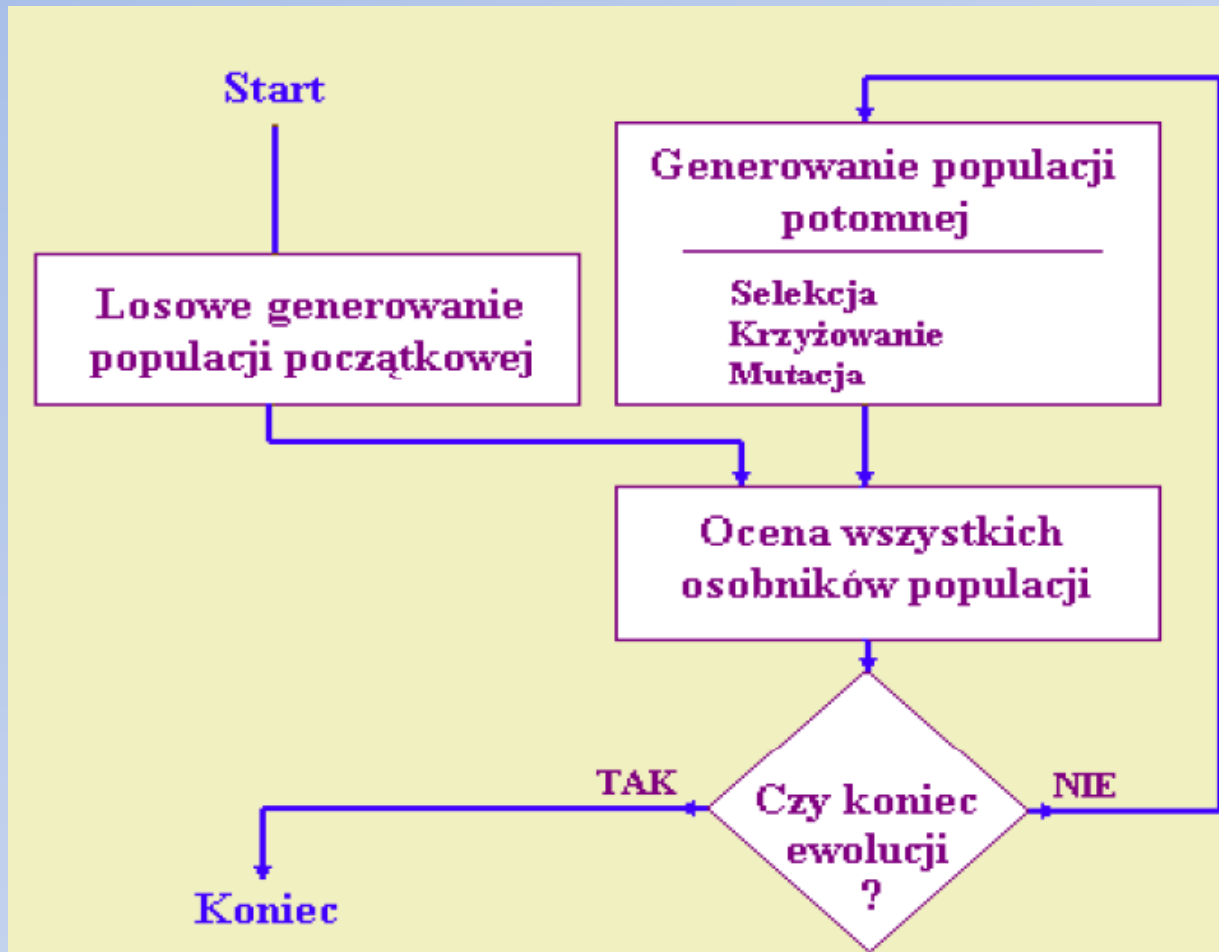
Algorytm genetyczny (genetic algorithm)- jest jedną z ewolucyjnych metod optymalizacji. Zalicza się go do klasy algorytmów heurystycznych.

Przeszukiwanie możliwych rozwiązań w celu znalezienia rozwiązania najlepszego lub potencjalnie najlepszego odbywa się za pomocą mechanizmów ewolucji oraz doboru naturalnego.

Zasada działania algorytmu genetycznego

1. Inicjujemy (najczęściej w sposób losowy) pewną początkową populację osobników
2. Poddajemy każdego z nich ocenie
3. Z populacji wybieramy osobniki najlepiej do tego przystosowane
4. Za pomocą operacji genetycznych (krzyżowanie oraz mutacja) tworzymy nowe pokolenie
5. Powrót do 2

Schemat działania GA



Potencjalne rozwiązania traktowane są jako **osobniki populacji**. Algorytm symuluje proces naturalnej selekcji poprzez ocenę **przystosowania** poszczególnych osobników, eliminację osobników słabszych i **krzyżowanie** ze sobą osobników najsilniejszych. Wynikiem działania algorytmu genetycznego jest populacja **najlepiej przystosowanych** osobników, wśród których może znajdować się najlepsze rozwiązanie. Jednocześnie najlepiej przystosowane osobniki nie muszą leżeć blisko siebie w przestrzeni rozwiązań.

Podstawowe pojęcia GA

Osobnik - pojedyncza propozycja rozwiązania problemu.

Populacja - zbiór osobników, na których operuje algorytm; odwzorowanie pewnych punktów z przestrzeni potencjalnych rozwiązań.

Chromosomy

Chromosom – reprezentacja potencjalnego rozwiązania podlegająca ocenie poprzez działania algorytmu genetycznego.

Chromosom może mieć różną postać w zależności od natury rozwiązywanego problemu.

Przykład:

Poszukiwanie ekstremum funkcji – postaciach chromosomów – *11000101, 01001111*.

Poszukiwanie ciągu znaków – chromosomy typu *hiasgfdo, qpom82ja, 7ama9g;1*.

Genotyp, fenotyp, geny

Genotyp to dziedziczna informacja, w którą wyposażony jest każdy osobnik.

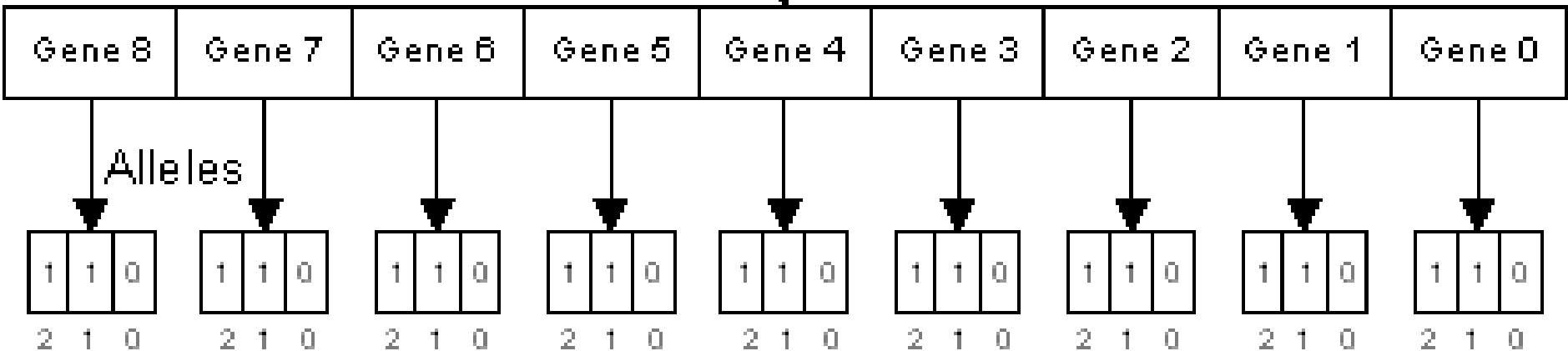
Fenotyp to zestaw konkretnych i dających się liczbowo zapisać cech generowanych na podstawie genotypu, które podlegają ocenie w środowisku.

Gen to najmniejszy element niosący informację genetyczną. Możliwymi wartościami genu są **allele**.

Individual



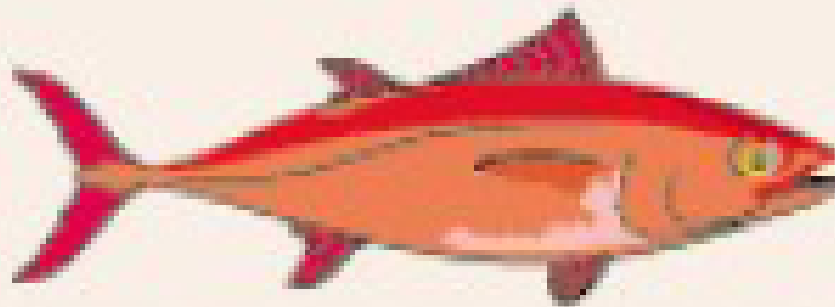
Genes



Przykład

kodowanie chromosomu

kształt	ogon	kolor	plełwy dolna	plełwy górna	zęby	wąsy
P - podłużny O - okrągły	M - mały D - duży	C - czerwony N - niebieski Z - zielony Ż - żółty	T - są N - brak	T - są N - brak	T - są N - brak	T - są N - brak



P M C T T N N

Kodowanie

Kodowanie – proces tworzenia fenotypu z genotypu; odwzorowanie rzeczywistych parametrów problemu za pomocą reprezentacji liczbowej.

Warianty ułożenia genów

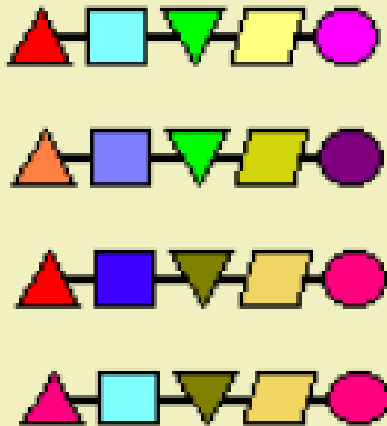
Klasyczny - geny na różnych pozycjach przechowują różne informacje. W wyniku krzyżowania geny nie zmieniają pozycji, lecz wartości. Wykorzystywany w problemach, gdzie chcemy dobrać optymalne cechy osobnika.

Permutacyjny - geny przechowują podobne informacje. W wyniku krzyżowania nie zmieniają wartości, lecz miejsce w chromosomie. Wykorzystywany w problemach kombinatorycznych, np. problemie komiwojażera.

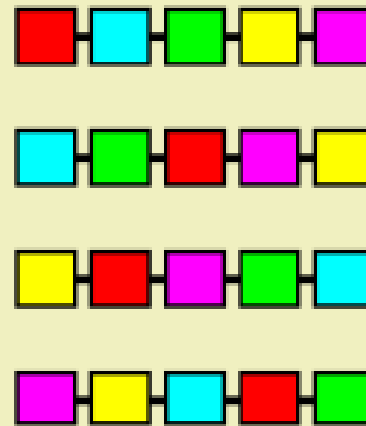
Drzewiasty - chromosom tworzy złożoną strukturę drzewiastą. W czasie krzyżowania przesunięciom ulegają całe gałęzie genów. Często geny mogą zmieniać także wartości. Wykorzystywany w programowaniu genetycznym oraz tam, gdzie ewolucji podlegają reguły matematyczne.

Warianty kodowania

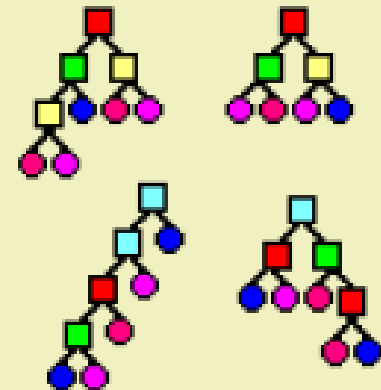
Kodowanie klasyczne



Kodowanie permutacyjne



Kodowanie drzewiaste



Funkcja przystosowania

Funkcja przystosowania – (ang. *fitness function*) – funkcja zwracająca liczbę, będącą oceną jakości przystosowania osobnika. Przystosowanie osobnika związane jest z jakością danego rozwiązania. Wartość ta jest generowana na podstawie fenotypu osobnika, a nie rzeczywistych cech z nim związanych.

Selekcja

W procesie selekcji wybierane są osobniki najlepiej przystosowane, które zostaną włączone do **grupy rozrodczej** i ich genotyp przetrwa do następnego pokolenia, tj. do następnej iteracji algorytmu. Istotne jest, aby właściwie dobrać stosunek wielkości tworzonej grupy do rozmiaru populacji. Zbyt małe stosunki (np. 1/1000) mogą doprowadzić do zaniku różnorodności genetycznej i defektów fenotypów, natomiast zbyt duże (np. 1/2) powodują wprowadzenie do rozrodu zbyt dużej liczby słabych genów, co również obniża jakość najlepszych osobników.

Metody selekcji

Metoda koła ruletki - wyobraźmy sobie koło ruletki, którego tarcza podzielona jest na pewną liczbę wycinków o różnych rozmiarach. Każdy osobnik otrzymuje obszar o rozmiarze wprost proporcjonalnym do jego jakości. Puszczamy koło ruletki w ruch, a po jego zatrzymaniu wskaźnik zatrzyma się na jakimś osobniku, który wchodzi do grupy rozrodczej. Im większy obszar dostał dany osobnik, tym większe prawdopodobieństwo jego wylosowania, dlatego też wielkość obszaru przydziela się przeważnie funkcją prawdopodobieństwa ($ZW = \langle 0, 1 \rangle$). Jeśli przez $F(i)$ oznaczymy wartość funkcji przystosowania osobnika i , to prawdopodobieństwo jego przetrwania wyliczymy ze wzoru:

$$p(i) = \frac{F(i)}{\sum_{i=1}^n F(i)}$$

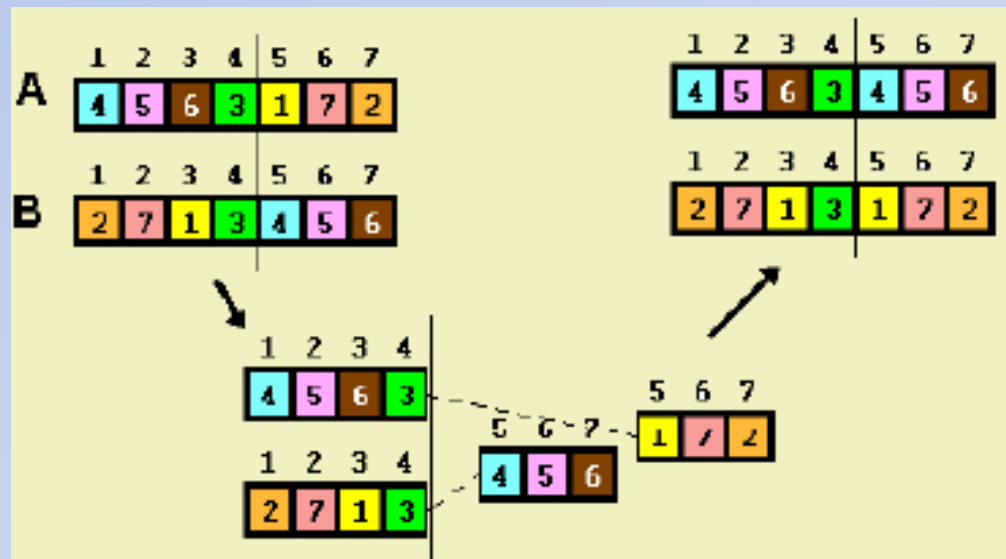
Metody selekcji

Selekcja rankingowa - osobniki populacji są sortowane według ich jakości: od najlepszego do najgorszego. Do dalszego rozrodu przechodzi tylko n najlepiej przystosowanych osobników. Metoda ta charakteryzuje się lepszym uwarunkowaniem niż metoda ruletki.

Selekcja turniejowa - populację dzieli się na szereg dowolnie licznych grup. Następnie z każdej z nich wybieramy osobnika o najlepszym przystosowaniu. Ten rodzaj selekcji także sprawdza się lepiej niż metoda ruletki.

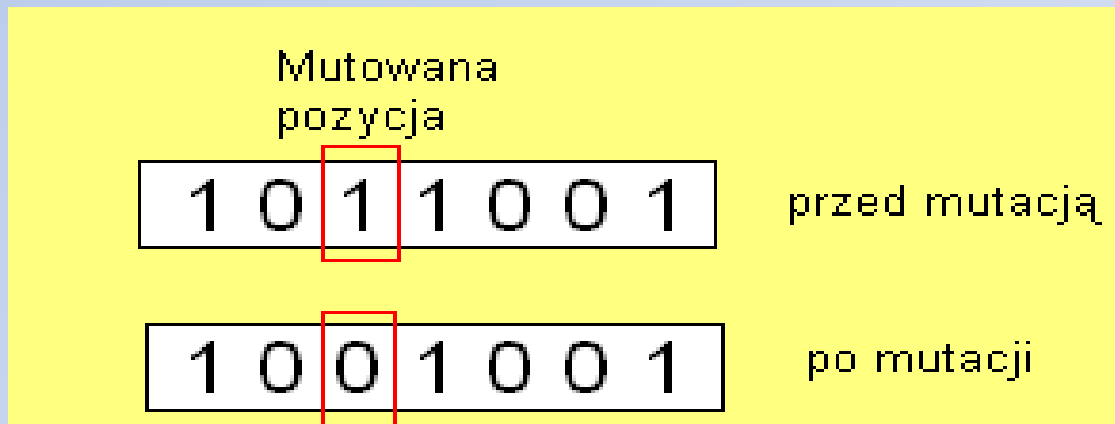
Krzyżowanie

Krzyżowanie (ang. *crossing-over*) – proces generowania nowej populacji z osobników, które weszły do grupy rozrodczej. Krzyżowanie odbywa się poprzez rozcięcie chromosomów rodziców w dowolnym punkcie. Następnie, połowa chromosomu przypada jednemu potomkowi, a połowa drugiemu.

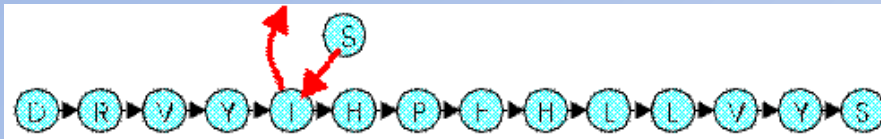


Mutacja

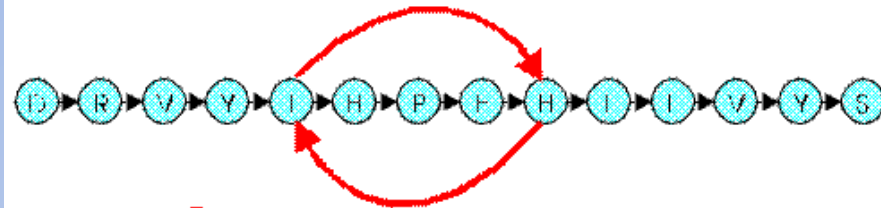
Mutacja to samoistna, losowa zmiana w genotypie. W przyrodzie powstaje w wyniku błędów reprodukcji lub uszkodzeń fizycznych i zwiększa różnorodność materiału genetycznego. Algorytm genetyczny symuluje zachodzenie mutacji. Prawdopodobieństwo jej zajścia to sprawa indywidualna każdego programu, lecz jego wartość powinna być bardzo niska, gdyż inaczej w wynikach pojawi się dużo losowości, a ewentualne korzystne zmiany nie będą miały szans na utrwalenie się w populacji.



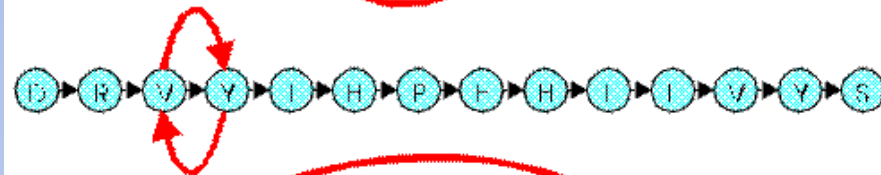
Metody mutacji



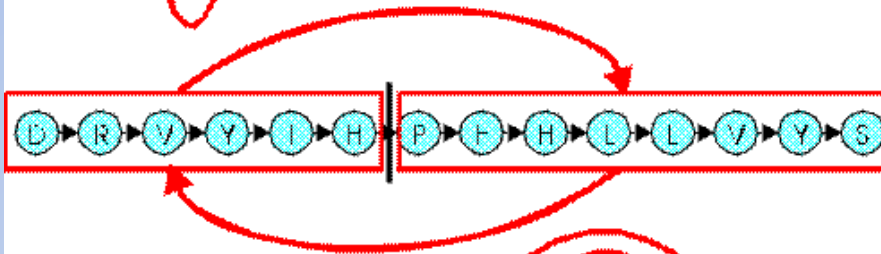
(a) Replacement



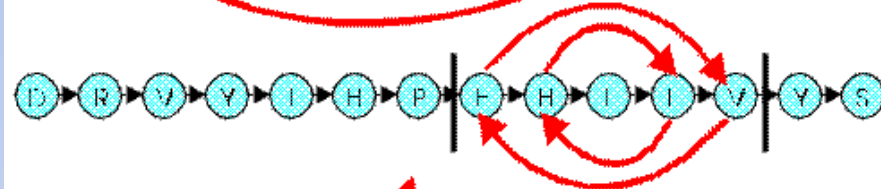
(b) Random swap



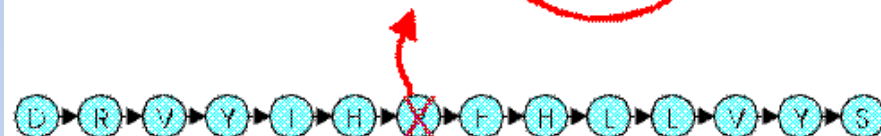
(c) Adjacent swap



(d) End-for-end swap

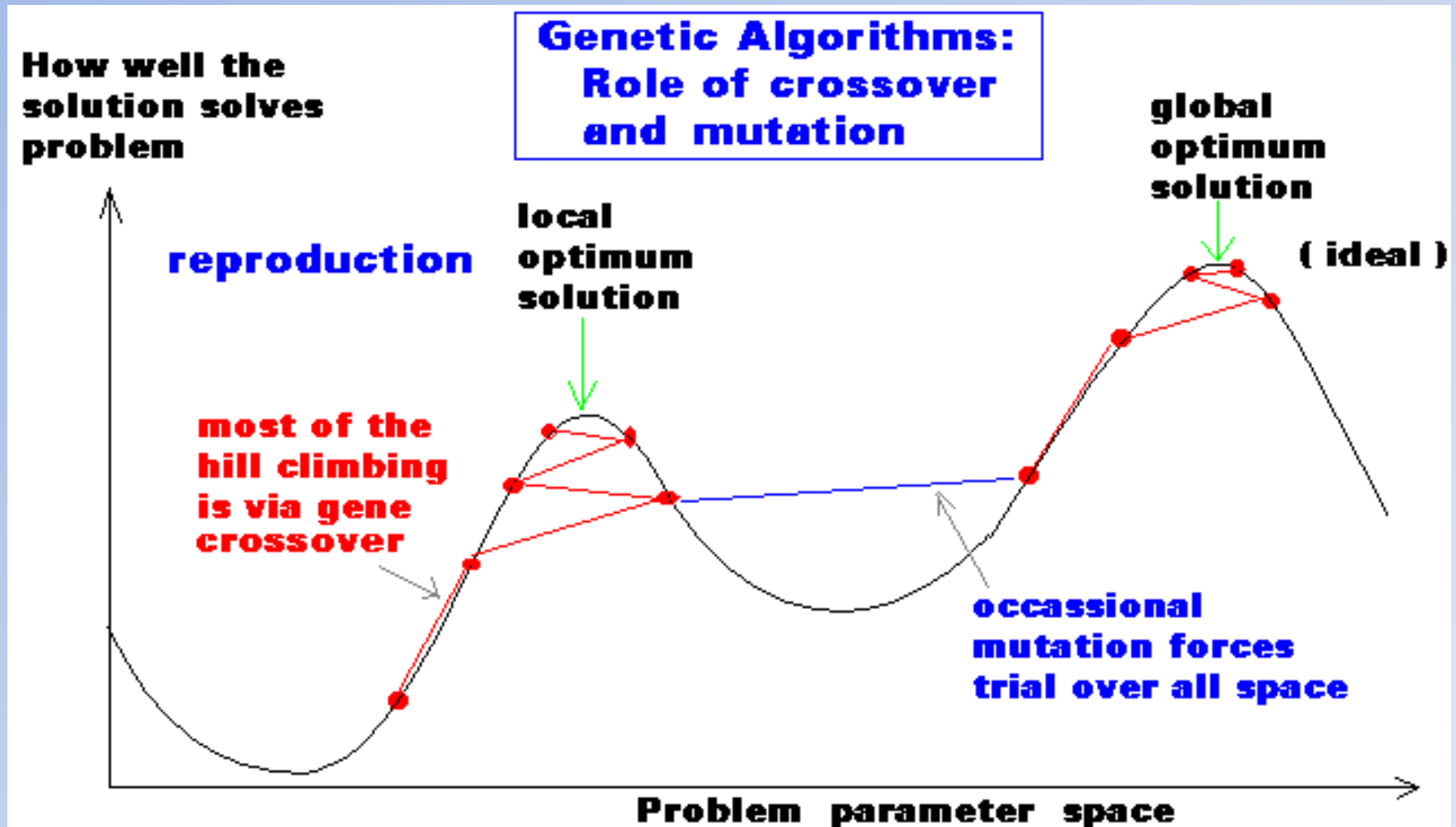


(e) Inversion



(f) Deletion

Znaczenie krzyżowania i mutacji



Demonstracje

<http://www.rennard.org/alife/english/gavgb.html>

<http://math.hws.edu/xJava/GA/>

http://userweb.elec.gla.ac.uk/y/yunli/ga_demo/

<http://www.blprnt.com/smartrockets/>