

Algorytmy genetyczne

(w materiałach wykorzystano wykłady prof.
Olivera de Wecka z MIT Open Courses)

Algorytmy heurystyczne

- Służą do znajdowania „wystarczająco dobrego” rozwiązania (nie mamy gwarancji, że rozwiązanie jest optymalne) w zadaniach, w których wykorzystanie innych metod jest problematyczne
- Często pozwalają na uniknięcie utknięcia w lokalnym ekstremum
- Niektóre umożliwiają na znajdowanie rozwiązań w zadaniach optymalizacyjnych, w których trudne, bądź wręcz niemożliwe jest wykorzystanie odpowiednich warunków koniecznych lub wystarczających (tam, gdzie np. nie można wykorzystać tzw. warunków Kuhna-Tuckera)

Przykłady heurystycznych algorytmów optymalizacji

- Genetic Algorithms (Algorytmy genetyczne)
 - Zainspirowane teorią ewolucji, genetyką i dobozem naturalnym
 - Maksymalizacja funkcji przystosowania
- Simulated Annealing (Symulowane wyżarzanie)
 - Zainspirowane procesami technologicznymi
 - Minimalizacja energii
- Particle Swarm Optimization
 - Zainspirowane zachowaniem rojów/stad (mrówek, pszczół, ptaków, wilków itp.
 - Maksymalizacja ilości zdobytego pożywienia

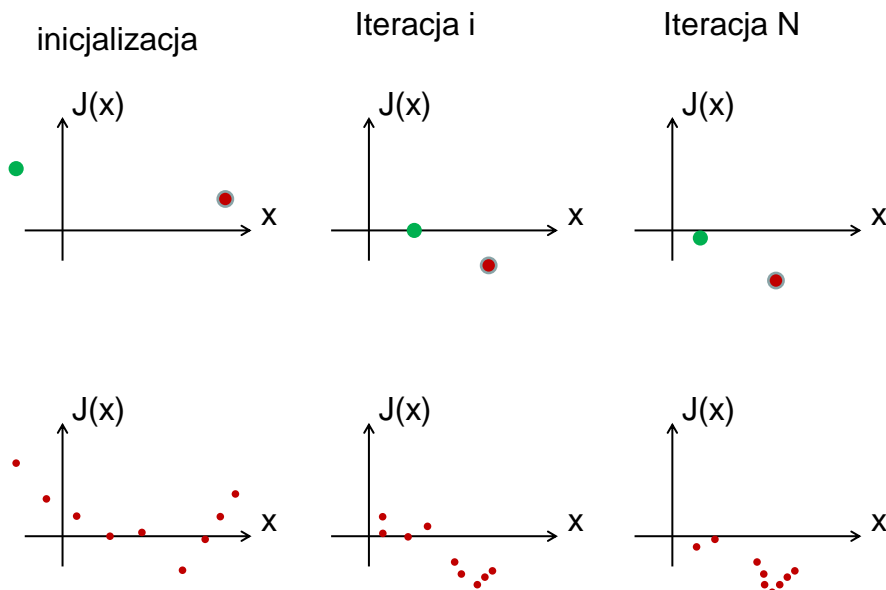
Algorytmy genetyczne – inspiracja

- U większości gatunków potencjał reprodukcyjny jest duży, ale tylko niewielki procent potomstwa dożywa wieku, w którym osiąga możliwości rozrodcze
- Cechy organizmu zakodowane są w genach
- O dostosowaniu się do środowiska decyduje fenotyp
- Dobór naturalny najlepiej przystosowanych osobników

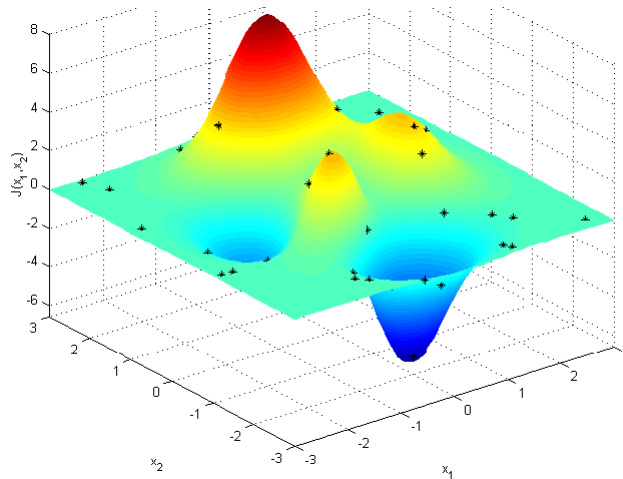
Algorytmy genetyczne - podstawowa idea

- Próbujemy odzwierciedlić proces doboru naturalnego, prowadzącego do najlepszego dostosowania
- Posługujemy się populacją rozwiązań danego problemu (jeden osobnik/chromosom – jedno rozwiązanie)
- Osobniki łączą się w pary, dając potomstwo, które dziedziczy ich cechy
- W populacji (w kolejnym pokoleniu) pozostają tylko te osobniki, których funkcja przystosowania ma największą wartość
- Funkcja przystosowania jest zdeterminowana przez cel zadania optymalizacji

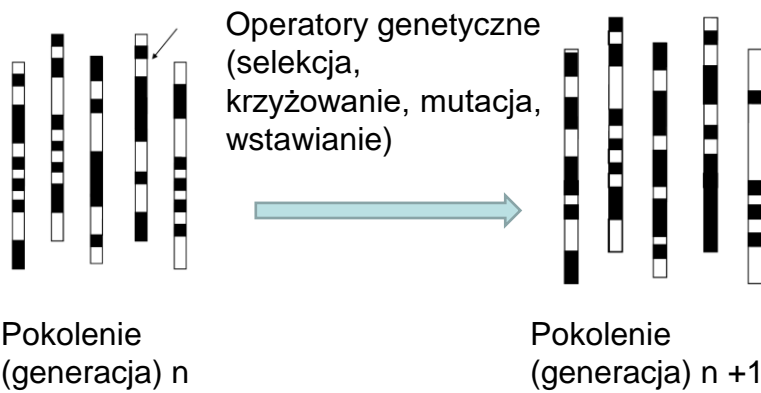
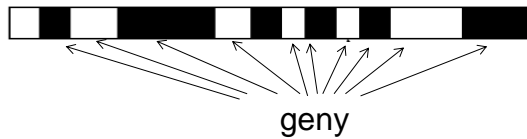
GA vs metody poszukiwań na kierunku (1D)



Populacja rozwiązań w problemie 2D



Pojedynczy osobnik (chromosom):



Podstawowy algorytm

1. Inicjalizacja (utworzenie początkowej populacji rozwiązań)
2. Wybór osobników „rodziców” do reprodukcji (selekcja)
3. Tworzenie osobników potomnych
 - Krzyżowanie (crossover)
 - Mutacja
4. Tworzenie kolejnej generacji (wstawianie osobników do nowej generacji - insercja)
5. Jeśli nie jest spełniony warunek stopu – powrót do 2.
6. Koniec

Implementacja

Należy

- Zdefiniować chromosom (geny) - kodowanie
- Zdefiniować funkcję przystosowania (jak genotyp przekłada się na fenotyp, a ten z kolei na przystosowanie) – dekodowanie i obliczanie funkcji przystosowania
- Określić rozmiar populacji
- Wybrać (zaprojektować) operator (algorytm) selekcji
- Wybrać (zaprojektować) operator (algorytm) krzyżowania
- Wybrać (zaprojektować) operator (algorytm) mutacji
- Określić kryterium stopu

Kodowanie/dekodowanie

Przykład 1 – problem plecakowy:

Wybór przedmiotów, które należy zapakować do plecaka, maksymalizujący ich sumaryczną wartość, przy ograniczeniach na pojemność/ sumaryczny ciężar plecaka

$$\max \sum_{j=1}^N c_j x_j \quad \text{przy ograniczeniu} \quad \sum_{j=1}^N w_j x_j \leq B$$

Gen – przypisany do przedmiotu:

- Wartość genu = 0 – przedmiot nie jest pakowany
- Wartość genu = 1 – przedmiot jest pakowany

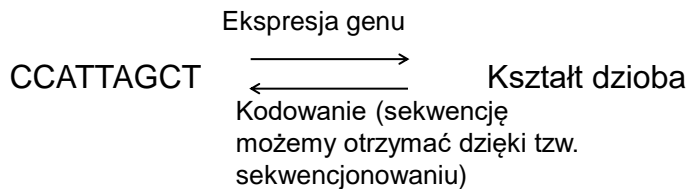
Funkcja przystosowania: $f = \sum_{j=1}^N c_j x_j$

Kodowanie - dekodowanie

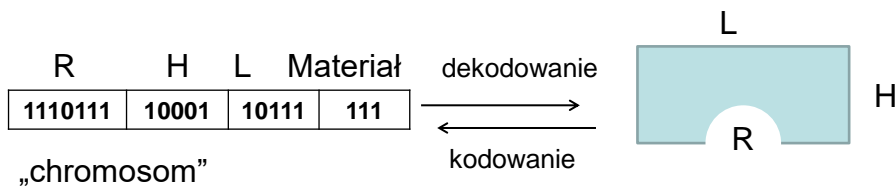
W biologii:

Genotyp
(Informacja
zapisana w DNA)

Fenotyp
(cecha, od której zależy
przystosowanie organizmu)



Problem optymalizacyjny



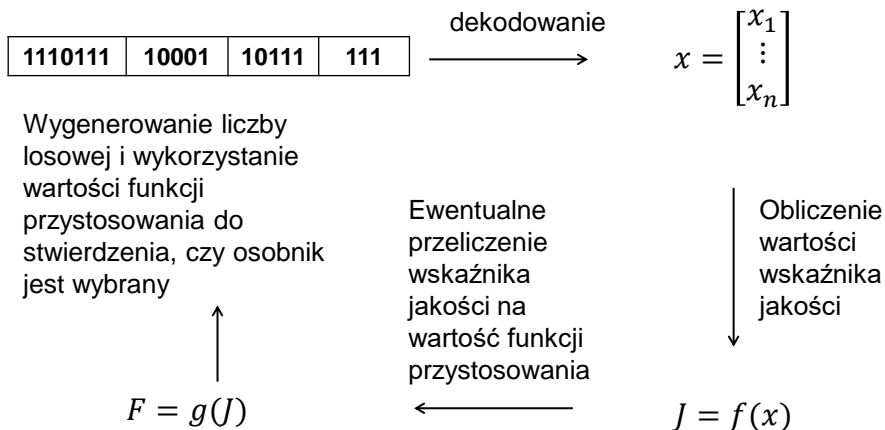
Kodowanie/dekodowanie - uwagi

Kodowanie

- Binarne (jeden gen – jeden bit)
- Liczby całkowite (właściwie dla określenia kolejności – np. w problemie komiwojażera)
- Elementy zdefiniowanego alfabetu
- Liczby rzeczywiste
- W zasadzie w pamięci wszystko jest zakodowane binarnie, ale rozróżnienie powyższych typów przydaje się do odpowiedniego dekodowania i wyboru operatorów krzyżowania
- W przypadku genów, reprezentujących wartości liczbowe, liczba bitów reprezentacji danej liczby z jednej strony determinuje dokładność, a z drugiej wymagany rozmiar populacji

Selekcja

- Wybór osobników (chromosomów) do reprodukcji – oparte na funkcji przystosowania
- Zwykle najbardziej złożony etap GA



Selekcja – metoda rankingowa

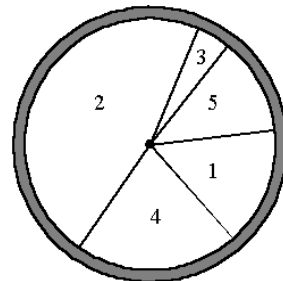
- Im wyższa pozycja w rankingu, tym większe prawdopodobieństwo wyboru osobnika
- Najprostsza implementacja – jeśli k oznacza pozycję osobnika w rankingu, to jest on losowany z prawdopodobieństwem np.. $1/(k+1)$
- Podstawowa wada - niewrażliwość na różnice w funkcji przystosowania pomiędzy kolejnymi osobnikami w rankingu

Selekcja – metoda ruletki

- Wirtualne koło ruletki
- Wycinki odpowiadają osobnikom
- Wielkość wycinka jest proporcjonalna do funkcji przystosowania

Prawdopodobieństwo selekcji k -tego osobnika: $f(k) / \sum f(i)$

	String s_i	Fitness $f(s_i)$	Relative Fitness
s_1	10110	2.23	0.14
s_2	11000	7.27	0.47
s_3	11110	1.05	0.07
s_4	01001	3.35	0.21
s_5	00110	1.69	0.11



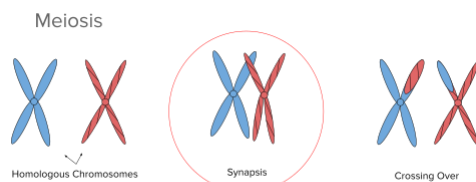
<https://neuron.eng.wayne.edu/tarek/MITbook/chap8/img00077.gif>

Selekcja – metoda turniejowa

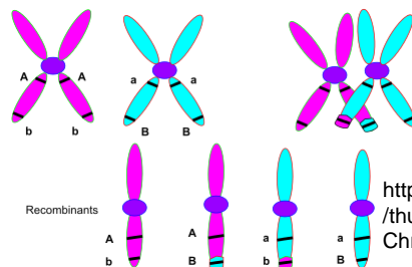
1. Losowy wybór dwóch rodziców
2. Zwycięzca wchodzi do zbioru rodziców
3. Jeśli założona pula rodziców nie została wypełniona- powrót do 1
4. W przeciwnym wypadku
 - Krzyżowanie
 - Mutacja
 - Wstawianie osobników do nowej populacji

Krzyżowanie

Tworzenie chromosomu jednego lub większej liczby potomków na podstawie chromosomów wybranych rodziców

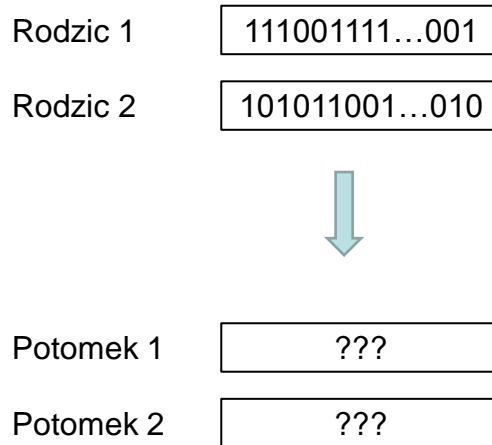


https://upload.wikimedia.org/wikipedia/commons/thumb/4/4c/Synapsis_during_Meiosis.svg/512px-Synapsis_during_Meiosis.svg.png



https://upload.wikimedia.org/wikipedia/commons/thumb/f/f0/Chromosomal_Crossover.svg/512px-Chromosomal_Crossover.svg.png

Krzyżowanie



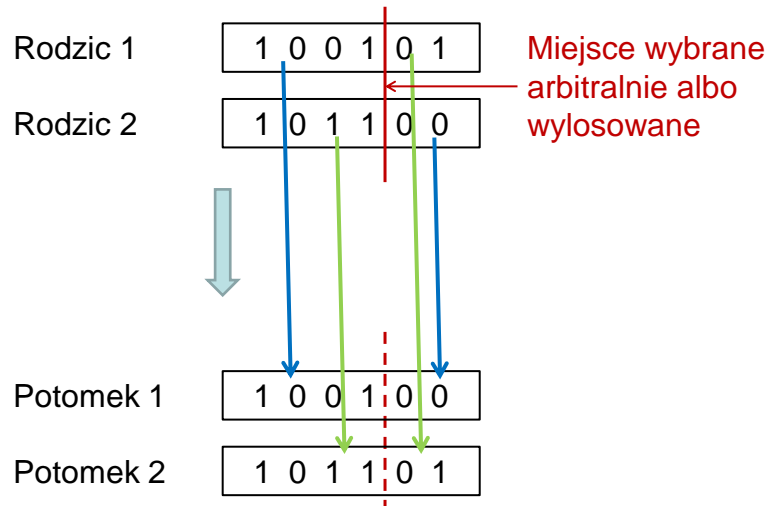
Krzyżowanie

Zależne od kodowania chromosomów i specyfiki problemu

Przykładowe operatory:

- rozcięcie dwóch chromosomów i stworzenie nowego poprzez sklejenie fragmentów pochodzących od rodziców (dla chromosomów z kodowaniem binarnym i całkowitoliczbowym)
 - Jednopunktowe
 - Wielopunktowe
- stosowanie operacji logicznych (dla chromosomów z kodowaniem binarnym),
- obliczenie wartości średniej genów (dla chromosomów z kodowaniem liczbami rzeczywistymi).

Krzyżowanie jednopunktowe

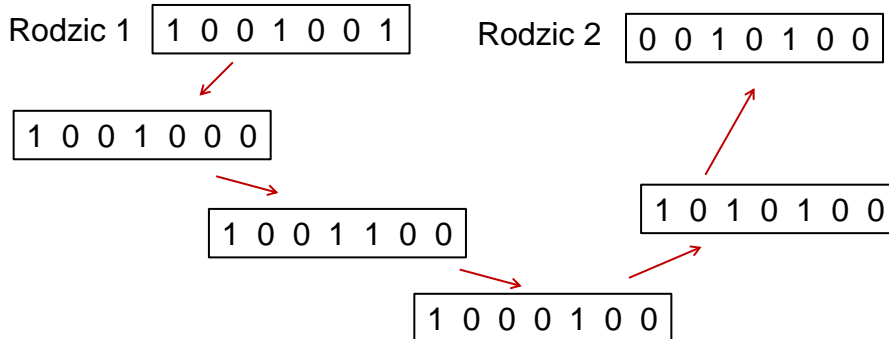


Krzyżowanie wielopunktowe



Krzyżowanie typu Path relinking

Utworzenie „ścieżki” wiodącej od jednego do drugiego rodzica



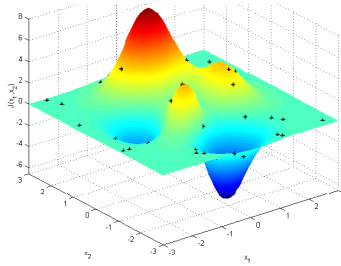
- Spośród tak utworzonych potomków wybierany jest ten o najlepszej wartości funkcji przystosowania
- Rozwiązania mają tendencję do interpolacji populacji początkowej

Krzyżowanie – inne metody

- Przy kodowaniu binarnym - stosowanie operacji logicznych na genach,
- Przy kodowaniu liczbami rzeczywistymi - obliczenie wartości średniej genów.

Mutacje

- Pozwalają na uniknięcie lokalnych ekstremów
- Mutacje punktowe lub permutacje
- Zbyt duży współczynnik mutacji – brak zbieżności algorytmu
- Zbyt mały współczynnik mutacji – niebezpieczeństwo utknięcia w lokalnym ekstremum



Tworzenie nowej populacji

Strategia globalne

- Zastąpienie całej populacji nowymi osobnikami (realizowane w różny sposób)

Strategia lokalna

- Zastąpienie pewnej wybranej liczby nowymi osobnikami (generowana jest mniejsza liczba nowych osobników – również różna realizacja), np..
 - Wybierz tylko dwóch rodziców, wygeneruj potomka i zastąp najslabszego osobnika
 - Pozostaw pewną liczbę najlepiej przystosowanych osobników z poprzedniej iteracji

Optymalizacja z ograniczeniami

Algorytmy genetyczne nie pozwalają na uwzględnianie ograniczeń w sposób jawny.

- Modyfikacja funkcji przystosowania – wprowadzenie funkcji kary
- Odpowiednia konstrukcja operatora genetycznego (np. odrzucanie osobników, którzy nie spełniają ograniczeń)
- Wprowadzenie operatora naprawy dla osobników, którzy nie spełniają ograniczeń
- Odpowiednie kodowanie

Inicjalizacja

Utworzenie początkowej populacji

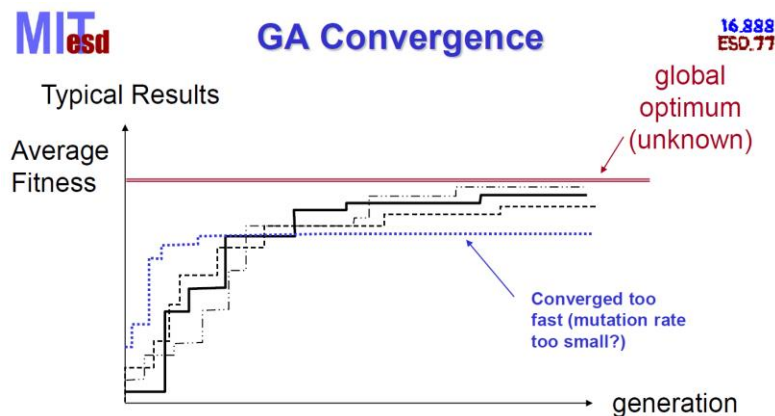
- Losowanie wartości genów, zwykle z rozkładu jednostajnego, tak by pokryć całą przestrzeń rozwiązań
 - Niekoniecznie dobre rozwiązanie
- Wielkość populacji
 - Osobników musi być odpowiednio dużo
 - Czasami można w podręcznikach znaleźć regułę, stanowiącą, że liczba osobników w populacji powinna być większa niż czterokrotna długość chromosomu (przy kodowaniu binarnym)

Kryterium stopu

- Założona liczba pokoleń
- Stabilizacja populacji
 - Różnica w sumarycznym przystosowaniu kolejnych populacji poniżej pewnego progu

$$|\sum_i J_i^{n+1} - \sum_i J_i^n| \leq \varepsilon$$
 - Zróżnicowanie w populacji (mierzone wariancją wartości funkcji przystosowania osobników) poniżej pewnego progu $\sigma^n \leq \delta$

Zbieżność algorytmu



Average performance of individuals in a population is expected to increase, as good individuals are preserved and bred and less fit individuals die out.

Podsumowanie

Algorytmy genetyczne w porównaniu do innych metod optymalizacji:

- Dokonują przeszukiwania przestrzeni rozwiązań równoległe, dla wielu punktów
- Wykorzystują probabilistyczne, nie deterministyczne, reguły przeszukiwania
- Nie wymagają dodatkowych informacji (np. o pochodnej funkcji celu), a jedynie podania postaci funkcji celu
- Wymagają opracowania specyficznego sposobu kodowania zmiennych i odpowiedniej implementacji i parametryzacji operatorów genetycznych