WYKŁAD 11

Algorytmy ewolucyjne

PLAN WYKŁADU

- Zasady konstruowania algorytmów ewolucyjnych
- Algorytmy genetyczne
  - schemat ogólny
  - operatory genetyczne
- Parę słów o teorii
- Strategie ewolucyjne

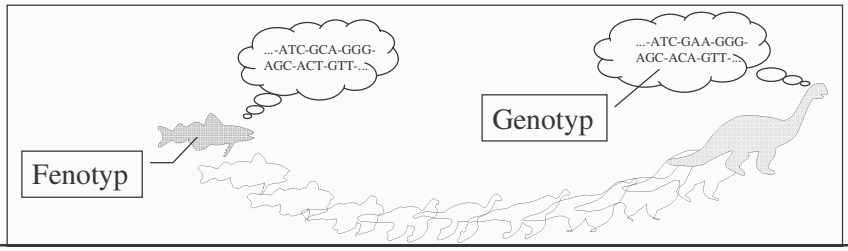
PODSTAWOWE POJĘCIA

Osobnik - podstawowa jednostka podlegająca ewolucji. Zakładamy zwykle, że ów osobnik przebywa w pewnym **środowisku**, do którego może być lepiej lub gorzej przystosowany. "Celem" ewolucji jest stworzenie osobnika możliwie dobrze przystosowanego do danego środowiska.

Fenotyp - ujawniające się "na zewnątrz" cechy danego osobnika.

Genotyp - "plan konstrukcyjny", kompletny i jednoznaczny opis osobnika zawarty w jego genach.

Populacja - zespół osobników zamieszkujących wspólne środowisko i konkurujących o jego zasoby.



PODSTAWOWE ZASADY

- **Genotyp danego osobnika w czasie jego życia nie ulega zmianie**, natomiast ulega on modyfikacjom podczas rozmnażania się. Zmiany te mogą wynikać albo z niewielkich, losowych **mutacji**, albo ze zmieszania (**skrzyżowania**) cech osobników rodzicielskich.
- **Zmiany w genotypie powodują zmiany fenotypu** osobników potomnych, co wpływa na stopień ich przystosowania do środowiska. Zmiany fenotypu (nabyte) nie podlegają dziedziczeniu w sensie genetycznym.
- **Zmiany w genotypie mają charakter przypadkowy**. Zmiany korzystne dla osobnika zdarzają się równie często, jak niekorzystne lub obojętne.
- **Osobniki są oceniane poprzez porównanie ich przystosowania do danego środowiska**. Te, które są lepiej przystosowane, mają większą szansę rozmnożyć się. Osobniki gorzej przystosowane przegrywają konkurencję o ograniczone zasoby środowiska i giną.

ZASTOSOWANIE:
PROBLEMY OPTYMALIZACYJNE

Osobnik (fenotyp)	⟷	Przykładowe rozwiązanie
Chromosom (genotyp)	⟷	Rozwiązanie zakodowane
Przystosowanie do środowiska	⟷	Optymalizowana funkcja

Zmianom (mutacja, krzyżowanie) podlega genotyp osobnika, podczas gdy selekcji poddawane są fenotypy. Istotą ewolucji jest połączenie zjawiska losowych, nieukierunkowanych zmian genotypu ze ściśle ukierunkowaną presją środowiska na fenotyp.

HISTORIA

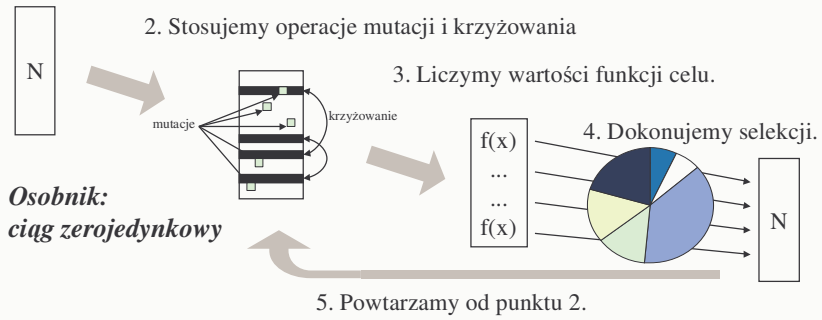
1958, 1964 (Friedberg, Fogel) - **programowanie ewolucyjne** automatów skończonych.
1965 (Bienert, Rechenberg, Schwefel) - **strategie ewolucyjne**, zastosowania praktyczne.
1975 (Holland) - **algorytmy genetyczne** i ich teoria.
Lata 80-te - liczne zastosowania algorytmów genetycznych.
Koniec lat 80-tych (Fogel) - współczesna wersja programowania ewolucyjnego.

- Z. Michalewicz. *Algorytmy genetyczne + struktury danych = programy ewolucyjne.*
- D.E. Goldberg. *Algorytmy genetyczne i ich zastosowania.* WNT, Warszawa 1995.
- J.R. Koza. *Genetic Programming: On the Programming of Computers by Means of the Natural Selection.* The MIT Press, 1992.

ALGORYTM GENETYCZNY

Cel: znaleźć maksimum funkcji f(x).
Założenie: funkcja ta jest dodatnia.

1. Tworzymy N osobników losowych.



5. Powtarzamy od punktu 2.

SCHEMAT DZIAŁANIA (1)

Etap wstępny: kodowanie problemu

Osobnik – ciąg binarny stałej długości.

Aby rozwiązać konkretne zadanie, musimy zakodować przestrzeń stanów (czyli wszystkie potencjalne rozwiązania) w języku binarnym.
Jeżeli zadanie polega na znalezieniu maksimum jakiejś funkcji, możemy owej funkcji użyć jako stopnia przystosowania osobnika do środowiska. Często musimy sami taką funkcję skonstruować.

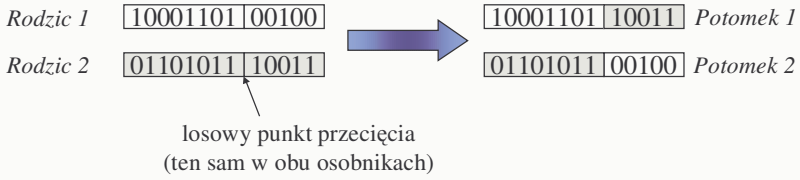
Drugi krok algorytmu: mutacje i krzyżowania

Mutacja: losujemy osobnika, następnie jeden z jego bitów. Zamieniamy wartość tego bitu na przeciwną. Mutacja dotyka średnio 0.1% bitów w populacji.

SCHEMAT DZIAŁANIA (2)

Krzyżowanie (*crossing-over*)

Łączymy osobniki w pary. Dla każdej pary ustalamy (w drodze losowania, prawdopodobieństwo rzędu 20-50%), czy dojdzie do ich skrzyżowania. Jeśli tak, losujemy miejsce (bit) w chromosomie jednego z rodziców, po czym zamieniamy miejscami fragmenty chromosomów poczynając od wylosowanego miejsca.



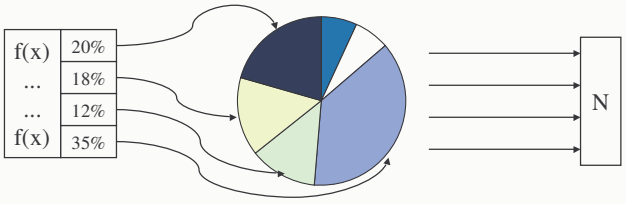
losowy punkt przecięcia
(ten sam w obu osobnikach)

SCHEMAT DZIAŁANIA (3)

Czwarty krok algorytmu: selekcja

Liczymy wartości funkcji celu osobników. Następnie, spośród N osobników populacji pośredniej losujemy N osobników populacji końcowej (z powtórzeniami), za pomocą algorytmu "koła ruletki":

1. Liczymy sumę wartości funkcji celu: $f_{sum} = f(x_1)+...+f(x_N)$.
2. Liczymy wkład każdego osobnika w sumę: $p(x_i) = f(x_i)/f_{sum}$
3. Dokonujemy N-krotnego losowania osobników zgodnie z rozkładem $p(x_i)$.



TEORIA

- **Schematy**: ciągi typu 100**1*1** (gwiazdka oznacza „cokolwiek”)
- **Twierdzenie o schematach**: jeśli schemat odpowiada osobnikom średnio lepszym niż inne, to jego reprezentanci są coraz liczniejsi (wzrost wykładniczy) w miarę trwania ewolucji
- **Ukryta równoległość**: w jednym kroku badamy niewiele osobników, jednak każdy z nich odpowiada wykładniczo wielu różnym schematom
- **Hipoteza o cegiełkach**: jeśli rozwiązanie zadania może być skonstruowane z silnych, krótkich schematów, wówczas algorytm genetyczny ma szansę łatwo je znaleźć

MOCNE STRONY

- **Metoda jest uniwersalna**. Aby tego samego programu użyć w innym problemie, przeważnie wystarczy zmienić funkcję celu.
- Algorytmy ewolucyjne potrafią poradzić sobie również tam, gdzie optymalizowana funkcja jest zaszumiona, zmienia się w czasie, ma wiele ekstremów lokalnych.
- Aby znaleźć rozwiązanie, nie musimy prawie nic wiedzieć o optymalizowanej funkcji (“**czarna skrzynka**”).
- **Metoda jest szybka**: znalezienie rozwiązania często jest możliwe po przejrzeniu zaskakująco niewielkiej części przestrzeni stanów.
- Ponieważ algorytm genetyczny jest algorytmem randomizowanym, możemy powtarzać obliczenia wielokrotnie w nadziei otrzymania lepszych wyników.

SŁABE STRONY

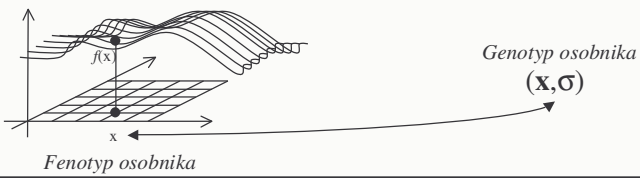
- Metoda jest uniwersalna, więc nie tak skuteczna, jak bywają algorytmy specjalizowane *(rada: stosować algorytmy hybrydowe).*
- Sukces jest możliwy wyłącznie przy prawidłowym zakodowaniu problemu i odpowiednim dobraniu funkcji celu. Niestety, nie ma jednoznacznej teorii mówiącej, jak to robić. Jest to często - podobnie jak dobór parametrów mutacji i krzyżowania - sprawa wyczucia i doświadczenia programisty *(rada: nabierać doświadczenia!).*
- Ponieważ algorytm genetyczny jest algorytmem randomizowanym, nigdy nie mamy pewności, że znaleźliśmy rozwiązanie optymalne *(rada: zadowolić się rozwiązaniem przybliżonym).*

STRATEGIE EWOLUCYJNE

Inna - obok algorytmów genetycznych - popularna metoda ewolucyjna. Główny obszar zastosowań: optymalizacja funkcji rzeczywistych wielowymiarowych.

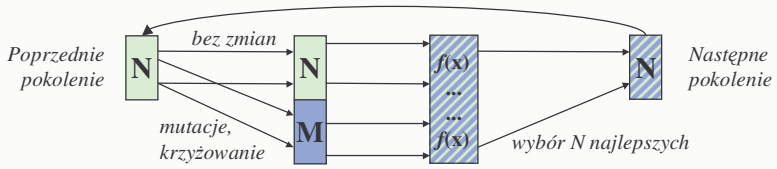
Zadanie: znaleźć maksimum funkcji $f : R^n \rightarrow R$ na danym przedziale $[a_1, b_1] \times ... \times [a_n, b_n]$.

Rozwiązanie: przez $\mathbf{x}=(x_1, ... x_n)$ oznaczmy przykładowe rozwiązanie (punkt z przestrzeni stanów, $x_i \in [a_i, b_i]$). Każdemu wektorowi **x** przyporządkujmy pomocniczy wektor wartości rzeczywistych dodatnich: $\sigma=(\sigma_1, ... \sigma_n)$. Para $(\mathbf{x},\sigma)$ będzie stanowiła kod genetyczny osobnika.



SCHEMAT DZIAŁANIA

1. Tworzymy losową populację złożoną z N osobników postaci (**x**,σ).
2. Dopisujemy do niej M osobników potomnych, tworzonych na podstawie losowo wybranych osobników z poprzedniego pokolenia za pomocą mutacji i krzyżowania.
3. Dla każdego osobnika w populacji pośredniej liczymy wartość optymalizowanej funkcji.
4. Spośród N+M osobników wybieramy N najlepszych względem *f(**x**)*. Te osobniki przeżyją i utworzą następne pokolenie.
5. Powtarzamy od punktu 2. z aktualną populacją N-osobnikową.



OPERATORY

Mutacji podlegają zwykle wszystkie osobniki dodawane do populacji pośredniej. Mutacja polega na zmianie parametrów (x,σ) zgodnie ze wzorami:

$$\sigma_i = \sigma_i \cdot e^{N(0,\Delta)}$$

$$x_i = x_i + N(0,\sigma_i)$$

gdzie N(0,σ) oznacza liczbę losową wygenerowaną zgodnie z rozkładem normalnym o wart. oczekiwanej 0 i odchyleniu standardowym σ.

Krzyżowaniu podlega część (np. 25%) osobników dodawanych do populacji pośredniej. Osobnik potomny składany jest z genów dwóch losowo wybranych osobników rodzicielskich.

$(x_1, x_2, x_3, x_4, x_5, \sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5)$

$(y_1, y_2, y_3, y_4, y_5, \tau_1, \tau_2, \tau_3, \tau_4, \tau_5)$

$\rightarrow (x_1, y_2, y_3, x_4, y_5, \sigma_1, \tau_2, \tau_3, \sigma_4, \tau_5)$

Uwaga: wartości x_i są dziedziczone zawsze razem z odpowiadającymi im wartościami pomocniczymi σ_i.