

Algorytmy ewolucyjne

Rafał Dreżewski

Akademia Górniczo–Hutnicza im. St. Staszica w Krakowie
Wydział Elektrotechniki, Automatyki, Informatyki i Elektroniki
Katedra Informatyki
drezew@agh.edu.pl

Podstawy biologiczne

- *Algorytmy ewolucyjne* to grupa technik znajdowania przybliżonych rozwiązań dla problemów optymalizacji i adaptacji, wzorowanych na mechanizmach ewolucji biologicznej.
- Teoria ewolucji dostarcza wyjaśnień dla dwóch grup zagadnień: powstawania różnych właściwości organizmów oraz powstawania różnych rodzajów organizmów (gatunków).
- Głównym procesem kształtującym właściwości organizmów jest dobór naturalny.

Podstawy biologiczne

- Istnienie doboru naturalnego jest faktem empirycznym i konsekwencją występowania pewnych podstawowych cech wszystkich żywych organizmów (śmiertelność, rozmnażanie się oraz dziedziczna zmienność cech wpływających na sposób ich funkcjonowania), jak również występowania ograniczonych zasobów środowiska, niezbędnych organizmom do przeżycia i reprodukcji.
- Współczesna, *syntetyczna teoria ewolucji*, która stanowi syntezę darwinizmu z osiągnięciami genetyki populacji, została opracowana w latach 20-tych i 30-tych XX wieku.

Pojęcia podstawowe

- Algorytm ewolucyjny przetwarza *populację osobników*, z których każdy jest punktem w przestrzeni potencjalnych rozwiązań pewnego problemu, dla którego zdefiniowana jest funkcja celu $f : D \rightarrow \mathbb{R}$.
- Funkcja celu może być zadana w postaci pewnego systemu świata rzeczywistego o dowolnej złożoności, symulacji komputerowej pewnego modelu lub też w postaci analitycznej.
- Każdy osobnik $a_i \in I$ (I jest przestrzenią osobników) posiada *genotyp* zawierający zestaw informacji niezbędnych do utworzenia *fenotypu* $\vec{x}^{a_i} \in D$.
- Genotyp osobnika składa się z *chromosomów*, z których przynajmniej jeden zawiera informacje kodujące fenotyp. Pozostałe chromosomy mogą zawierać pewne parametry istotne dla działania samego algorytmu ewolucyjnego.

Pojęcia podstawowe

- Chromosom składa się z kolei z *genów* będących jednostkami elementarnymi genotypu osobnika.
- W *środowisku*, w którym działa algorytm ewolucyjny, zdefiniowana jest tzw. *funkcja przystosowania* $\varphi : I \rightarrow \mathbb{R}$.
- Funkcja przystosowania, w ogólnym przypadku, nie musi być identyczna z funkcją celu f , jednakże funkcja celu jest zawsze jej częścią składową.
- Funkcja przystosowania (φ) jest na ogół złożeniem funkcji celu (f) i funkcji dekodującej fenotyp osobnika ($fen : I \rightarrow D$; $fen' : D \rightarrow I$ jest funkcją kodującą fenotyp osobnika): $\varphi = f \circ fen$.

Pojęcia podstawowe

- Populacja początkowa jest generowana losowo (lub z uwzględnieniem wiedzy o problemie) i stopniowo ewoluuje w kierunku coraz „lepszych” (z punktu widzenia rozwiązywanego problemu) obszarów przestrzeni poszukiwań dzięki zastosowaniu probabilistycznego (lub deterministycznego) mechanizmu *selekcji* oraz operatorów *rekombinacji* i/lub *mutacji*.

Pojęcia podstawowe

- Procesem *selekcji* określa się w literaturze łącznie proces *reprodukcji*, w którym powielane są losowo wybrane z populacji osobniki oraz proces *sukcesji*, w trakcie którego z *populacji potomnej* wybierane są osobniki tworzące *populację bazową* kolejnego pokolenia
- Proces selekcji faworyzuje osobniki o lepszym przystosowaniu, dzięki czemu mają one większe szanse na reprodukcję i wprowadzenie do kolejnego pokolenia swojego potomstwa niż osobniki gorzej przystosowane.

Pojęcia podstawowe

- Mechanizm *rekombinacji* umożliwia wymianę informacji pomiędzy różnymi osobnikami populacji poprzez „wymieszanie” odpowiednich chromosomów pochodzących od różnych rodziców w trakcie przekazywania materiału genetycznego potomstwu.
- Mechanizm *mutacji* odpowiedzialny jest za wprowadzanie innowacji. Jego działanie polega na losowym zaburzeniu genotypu potomka, przy czym najczęściej niewielkie zaburzenia są bardziej prawdopodobne od dużych.

Zastosowania

- Algorytmy ewolucyjne znalazły zastosowanie praktyczne w wielu różnych dziedzinach, na przykład:
 - w oprogramowaniu typu CAD (projektowanie kształtu komory silnika odrzutowego, rozłożenia elementów na płycie krzemu, projektowanie anten);
 - w badaniach operacyjnych (harmonogramowanie zadań, planowanie transportu);
 - we wspomaganiu nawigacji;
 - w planowaniu tras robotów;
 - w narzędziach wspomagania decyzji ekonomicznych;
 - sztucznej inteligencji i maszynowym uczeniu (poszukiwanie optymalnych reguł klasyfikacji);
 - symulacjach z dziedziny sztucznego życia.

Schemat działania algorytmu ewolucyjnego

```
t := 0;  
wygenerowanie  $A(t)$ ;  
ocena  $A(t)$ ;  
while not warunek stopu do  
  begin  
     $A^1(t)$  := reprodukcja  $A(t)$ ;  
     $A^2(t)$  := rekombinacja  $A^1(t)$ ;  
     $A^3(t)$  := mutacja  $A^2(t)$ ;  
    ocena  $A^3(t)$ ;  
     $A(t+1)$  := sukcesja  $(A^3(t) \cup A^4(t))$ ;  
    t := t + 1;  
  end
```

Podstawowe rodzaje algorytmów ewolucyjnych

- *Strategie ewolucyjne* (ang. *evolution strategies* — *ES*) opracowane przez I. Rechenberg'a oraz H.-P. Schwefel'a.
- *Programowanie ewolucyjne* (ang. *evolution programming* — *EP*) opracowane przez L. J. Fogel'a.
- *Algorytmy genetyczne* (ang. *genetic algorithms* — *GA*) opracowane przez J. H. Holland'a.

Strategie ewolucyjne

- Strategie ewolucyjne są aktualnie najczęściej wykorzystywane w problemach optymalizacji ciągłej, gdzie funkcja celu jest określona następująco:

$$f : \mathbb{R}^n \supseteq D \rightarrow \mathbb{R} \quad (1)$$

- Dla strategii ewolucyjnych charakterystyczna jest reprezentacja zmienno pozycyjna, deterministyczny mechanizm sukcesji, nadawanie głównego znaczenia operatorowi mutacji oraz mechanizm autoadaptacji zasięgu mutacji.

Strategie ewolucyjne

- W uproszczonym przypadku (gdy operator mutacji nie wykorzystuje macierzy kowariancji) osobnik $a = \langle \vec{x}, \vec{\sigma} \rangle$ składa się z wektora wartości zmiennych niezależnych $\vec{x} \in D$ oraz wektora odchyleń standardowych wykorzystywanych podczas mutacji $\vec{\sigma}$, co ma na celu umożliwienie autoadaptacji zasięgu mutacji.

Strategie ewolucyjne

Schemat działania współczesnych strategii ewolucyjnych można opisać następująco:

- W strategii (μ, λ) , w pokoleniu t , tworzonych jest λ potomków z μ rodziców ($1 \leq \mu < \lambda$) z wykorzystaniem operatorów reprodukcji, rekombinacji i mutacji. Następnie spośród λ potomków (czyli ze zbioru $A^3(t)$) wybieranych jest μ najlepszych osobników, z których zostaje utworzona populacja bazowa pokolenia $t + 1$.
- W strategii $(\mu + \lambda)$, w pokoleniu t , również tworzonych jest λ potomków z μ rodziców z wykorzystaniem operatorów reprodukcji, rekombinacji i mutacji ($1 \leq \mu \leq \lambda$). Następnie, spośród $\mu + \lambda$ osobników (czyli ze zbioru $A^3(t) \cup A(t)$) wybieranych jest μ najlepszych osobników w celu utworzenia populacji bazowej pokolenia $t + 1$.

Strategie ewolucyjne

W jednej z wersji krzyżowania uśredniającego, i -ta wartość wektora \vec{x}^{a_l} potomka a_l rodziców a_j i a_k jest określona wzorem:

$$x_i^{a_l} = \xi_{U(0;1),i} x_i^{a_j} + (1 - \xi_{U(0;1),i}) x_i^{a_k} \quad (2)$$

gdzie $x_i^{a_j}$ jest i -tą wartością wektora \vec{x}^{a_j} osobnika a_j , $x_i^{a_k}$ jest i -tą wartością wektora \vec{x}^{a_k} osobnika a_k , $\xi_{U(0;1),i}$ jest realizacją zmiennej losowej o rozkładzie jednostajnym na odcinku $(0; 1)$ dla i -tej wartości wektora \vec{x}^{a_l} . Operator ten jest stosowany również w przypadku wektora $\vec{\sigma}$:

$$\sigma_i^{a_l} = \xi_{U(0;1),i} \sigma_i^{a_j} + (1 - \xi_{U(0;1),i}) \sigma_i^{a_k} \quad (3)$$

Strategie ewolucyjne

W jednej z wersji operatora mutacji z mechanizmem autoadaptacji zasięgu, wzór określający nową i -tą wartość wektora odchyłeń standardowych $\vec{\sigma}$ osobnika a jest następujący:

$$\sigma'_i = \sigma_i \exp(\tau_0 \xi_{N(0,1)} + \tau \xi_{N(0,1),i}) \quad (4)$$

gdzie $\xi_{N(0,1)}$ jest realizacją zmiennej losowej o standaryzowanym rozkładzie normalnym identyczną dla wszystkich elementów wektora $\vec{\sigma}$, natomiast $\xi_{N(0,1),i}$ jest realizacją zmiennej losowej o standaryzowanym rozkładzie normalnym dla i -tej wartości wektora $\vec{\sigma}$.

Strategie ewolucyjne

Zalecane wartości τ_0 oraz τ wynoszą:

$$\tau_0 = \frac{k_1}{\sqrt{2n}} \quad (5)$$

$$\tau = \frac{k_2}{\sqrt{2\sqrt{n}}} \quad (6)$$

gdzie k_1 oraz k_2 są pewnymi stałymi.

Strategie ewolucyjne

W drugiej kolejności wykonywana jest mutacja wektora wartości zmiennych niezależnych \vec{x} osobnika a , zgodnie z formułą:

$$x'_i = x_i + \sigma'_i \xi_{N(0,1)} \quad (7)$$

Strategie ewolucyjne

$t := 0;$

wygenerowanie $A(t) = \{a_1(t), \dots, a_\mu(t)\};$

ocena $A(t);$

while not *warunek stopu* **do**

begin

$A^1(t) :=$ wylosowanie λ osobników z $A(t);$

$A^2(t) :=$ rekombinacja $A^1(t);$

$A^3(t) :=$ mutacja $A^2(t);$

ocena $A^3(t);$

if (μ, λ) -ES **then**

$A(t+1) := \mu$ najlepszych osobników z $A^3(t)$

else

$A(t+1) := \mu$ najlepszych osobników
z $(A^3(t) \cup A(t))$

$t := t + 1;$

end

Programowanie ewolucyjne

- Programowanie ewolucyjne w swej oryginalnej formie było próbą uzyskania inteligentnego zachowania na drodze symulowanej ewolucji, natomiast współcześnie jest ono głównie stosowane w problemach optymalizacji ciągłej.
- W programowaniu ewolucyjnym stosowana jest również reprezentacja zmiennopozycyjna, dostosowany do niej operator mutacji oraz mechanizm autoadaptacji zasięgu operatora mutacji.
- Osobnik $a = \langle \vec{x}, \vec{\sigma} \rangle$ składa się z wektora wartości zmiennych niezależnych $\vec{x} \in D$ oraz wektora odchylen standardowych wykorzystywanych podczas mutacji $\vec{\sigma}$.

Programowanie ewolucyjne

- Osobnik $a = \langle \vec{x}, \vec{\sigma} \rangle$ składa się z wektora wartości zmiennych niezależnych $\vec{x} \in D$ oraz wektora odchyleń standardowych wykorzystywanych podczas mutacji $\vec{\sigma}$.
- W jednej z wykorzystywanych w programowaniu ewolucyjnym wersji operatora mutacji, wzór określający nową i -tą wartość wektora \vec{x} osobnika a jest następujący:

$$x'_i = x_i + \sigma_i \xi_{N(0,1)} \quad (8)$$

gdzie $\xi_{N(0,1)}$ jest realizacją zmiennej losowej o standaryzowanym rozkładzie normalnym identyczną dla wszystkich elementów wektora \vec{x} .

Programowanie ewolucyjne

- W drugiej kolejności wykonywana jest mutacja wektora odchyłeń standardowych $\vec{\sigma}$ osobnika a , zgodnie z formułą:

$$\sigma'_i = \sigma_i + \kappa \sigma_i \xi_{N(0,1)} \quad (9)$$

gdzie κ jest parametrem sterującym intensywnością zmian σ_i . Jeżeli w rezultacie zastosowania powyższego wzoru $\sigma'_i < 0$ wtedy przyjmuje się, że $\sigma'_i = \varepsilon$, przy czym wartość $\varepsilon > 0$ jest ustalona i dowolnie bliska zera.

Programowanie ewolucyjne

```
t := 0;
wygenerowanie  $A(t) = \{a_1(t), \dots, a_\mu(t)\}$ ;
ocena  $A(t)$ ;
while not warunek stopu do
begin
     $A^1(t) := A(t)$ ;
     $A^3(t) := \text{mutacja } A^1(t)$ ;
    ocena  $A^3(t)$ ;
     $A(t+1) := \mu$  najlepszych w turnieju
                           osobników z  $(A^3(t) \cup A(t))$ ;
    t := t + 1;
end
```

Algorytmy genetyczne

- Główne cechy odróżniające zaproponowany przez J. Holland'a tzw. *prosty algorytm genetyczny* od innych rodzajów algorytmów ewolucyjnych to: reprezentacja binarna, probabilistyczny mechanizm reprodukcji oraz nadawanie głównego znaczenia operatorowi rekombinacji.
- W przypadku reprezentacji binarnej chromosom osobnika $a = \langle (b_1, \dots, b_m) \rangle$ jest m elementowym wektorem genów, z których każdy należy do zbioru $\{0, 1\}$, natomiast przestrzeń osobników jest określona następująco:
 $I = \{0, 1\}^m$.

Algorytmy genetyczne

- W celu zastosowania prostego algorytmu genetycznego do problemów optymalizacji ciągłej, w których funkcja celu ma postać $f : \mathbb{R}^n \supseteq D \rightarrow \mathbb{R}$ konieczne jest podzielenie chromosomu na n logicznych segmentów o (na ogół) równej długości m_x , z których każdy jest interpretowany jako kod binarny odpowiedniej wartości wektora zmiennych niezależnych \vec{x} (fenotypu osobnika a).
- Osobnik jest w takim przypadku zdefiniowany następująco:
$$a = \langle (b_{11}, \dots, b_{1m_x}, b_{21}, \dots, b_{2m_x}, \dots, b_{n1}, \dots, b_{nm_x}) \rangle \in$$
$$I = \{0, 1\}^m, \text{ gdzie } m = n \cdot m_x.$$
- Funkcja przystosowania jest zazwyczaj złożeniem funkcji celu i funkcji dekodującej fenotyp osobnika $\varphi = f \circ fen$, gdzie $fen : \{0, 1\}^m \rightarrow D$ jest funkcją dekodującą fenotyp osobnika, natomiast $fen' : D \rightarrow \{0, 1\}^m$ jest funkcją kodującą fenotyp.

Algorytmy genetyczne

- W algorytmach genetycznych wykorzystywane są na ogół probabilistyczne mechanizmy reprodukcji.
- Jednym z takich mechanizmów jest zaproponowana przez J. Holland'a tzw. *reprodukcja proporcjonalna*, w której osobniki o lepszym przystosowaniu mają większe szanse na reprodukcję.
- W tym celu określa się zmienną losową na zbiorze $A(t)$ (populacji pokolenia t) o następującym rozkładzie:

$$p_{rep}(a_i) = \frac{\varphi(a_i)}{\sum_{j=1}^{\mu} \varphi(a_j)}, \quad \forall j \ a_j, a_i \in A(t) \quad (10)$$

gdzie a_i jest osobnikiem, natomiast $\varphi(a_i)$ jest wartością funkcji przystosowania dla osobnika a_i . Następnie μ razy wykonywana jest realizacja tej zmiennej losowej i wylosowane osobniki są kopiowane do zbioru $A^1(t)$.

Algorytmy genetyczne

- W przypadku mechanizmu reprodukcji proporcjonalnej, funkcja przystosowania powinna przyjmować wyłącznie wartości większe od zera oraz zapewniać, że im dany osobnik stanowi lepszą propozycję rozwiązania danego problemu, tym wartość jego przystosowania jest większa.
- W tym celu wprowadza się specjalną funkcję skalującą $\delta : \mathbb{R} \rightarrow \mathbb{R}$, natomiast funkcja przystosowania jest wtedy zdefiniowana następująco:

$$\varphi = \delta \circ f \circ fen \quad (11)$$

Funkcja skalująca δ zapewnia, że funkcja przystosowania spełnia wspomniane powyżej kryteria.

- Wśród innych stosowanych w algorytmach genetycznych mechanizmów reprodukcji wymienić można *reprodukcję rangową (rankingową)* oraz *reprodukcję turniejową*.

Algorytmy genetyczne

- W algorytmach genetycznych operatorem o głównym znaczeniu jest operator rekombinacji. Parametr p_{rec} określa prawdopodobieństwo zastosowania operatora rekombinacji w trakcie tworzenia osobnika potomnego.
- W prostym algorytmie genetycznym jako operator rekombinacji wykorzystywane jest tzw. *krzyżowanie jednopunktowe*. Polega ono na rozcięciu chromosomów osobników rodzicielskich, $a_i = \langle (b_1^{a_i}, \dots, b_m^{a_i}) \rangle$ oraz $a_j = \langle (b_1^{a_j}, \dots, b_m^{a_j}) \rangle$, na dwa fragmenty, a następnie zamianie odpowiadających sobie fragmentów i utworzeniu osobników potomnych a_k oraz a_l :

$$a_k = \langle (b_1^{a_i}, \dots, b_{z-1}^{a_i}, b_z^{a_j}, b_{z+1}^{a_j}, \dots, b_m^{a_j}) \rangle \quad (12a)$$

$$a_l = \langle (b_1^{a_j}, \dots, b_{z-1}^{a_j}, b_z^{a_i}, b_{z+1}^{a_i}, \dots, b_m^{a_i}) \rangle \quad (12b)$$

Miejsce rozcięcia $z \in \{1, \dots, m\}$ jest wybierane losowo z rozkładem równomiernym.

Algorytmy genetyczne

- Krzyżowanie jednopunktowe można łatwo uogólnić do *krzyżowania p -punktowego*, w którym wybieranych jest p miejsc rozcięcia, a następnie wykonywana jest zamiana co drugiego odpowiadającego sobie fragmentu chromosomów.
- *Krzyżowanie równomierne* polega z kolei na losowym podejmowaniu decyzji o zamianie każdego odpowiadającego sobie genu z chromosomów rodziców w trakcie tworzenia chromosomów potomków.

Algorytmy genetyczne

- Mutacja jest traktowana w algorytmach genetycznych jako operator o zdecydowanie mniejszym znaczeniu niż rekombinacja. Mutacja jest wykonywana oddzielnie dla każdego genu i polega na zmianie wartości genu na przeciwną z prawdopodobieństwem p_{mut} .
- W przypadku osobnika $a = \langle (b_1, \dots, b_m) \rangle$ działanie operatora mutacji jest określone następująco:

$$\forall i \in \{1, \dots, m\} \quad b'_i = \begin{cases} b_i & \text{jeżeli } \xi_{U[0;1],i} > p_{mut} \\ 1 - b_i & \text{jeżeli } \xi_{U[0;1],i} \leq p_{mut} \end{cases} \quad (13)$$

gdzie $\xi_{U[0;1],i}$ jest realizacją zmiennej losowej o rozkładzie jednostajnym na odcinku $[0; 1]$ dla genu b_i .

Algorytmy genetyczne

- W prostym algorytmie genetycznym najczęściej stosowana jest *sukcesja z całkowitym zastępowaniem* (tzw. *sukcesja trywialna*). Nową populacją bazową staje się tutaj populacja potomna ($A(t+1) = A^3(t)$).
- Wśród innych mechanizmów sukcesji wymienić można *sukcesję elitarną*, w której nowa populacja bazowa tworzona jest z μ najlepszych osobników wybranych ze zbioru $A^3(t) \cup A^4(t)$, gdzie $A^3(t)$ jest populacją potomną, natomiast $A^4(t)$ to zbiór n najlepszych osobników ze starej populacji bazowej.

Algorytmy genetyczne

```
 $t := 0;$   
wygenerowanie  $A(t) = \{a_1(t), \dots, a_\mu(t)\};$   
ocena  $A(t);$   
while not warunek stopu do  
  begin  
     $A^1(t) :=$  reprodukcja  $A(t);$   
     $A^2(t) :=$  rekombinacja  $A^1(t);$   
     $A^3(t) :=$  mutacja  $A^2(t);$   
    ocena  $A^3(t);$   
     $A(t+1) := A^3(t);$   
     $t := t + 1;$   
  end
```
