Algorytmy genetyczne (AG)

- 1. Wprowadzenie do AG
 - a) ewolucja darwinowska
 - b) podstawowe definicje
 - c) operatory genetyczne
 - d) konstruowanie AG
 - e) standardowy AG
 - f) przykład rozwiązania
 - g) naprawdę bardzo, bardzo nieco historii
 - h) literatura (kilka pozycji)
- 2. Algorytmy ewolucyjne (ledwie wspomnienie)
- 3. Standardowy Algorytm Genetyczny (SAG)
- 4. AG a TSP

- 1. Wprowadzenie do AG
- a) ewolucja darwinowska
 - 1. każde środowisko ma ograniczone zasoby określa to maksymalną liczbę osobników (populacja) mogących utrzymać się w nim przy życiu,
 - 2. liczba osobników zmienia się w czasie najczęściej rośnie; osobniki rozmażają się → populacja się powiększa,
 - 3. z 1. i 2. wynika, że liczba osobników musi być regulowana środowisko określa warunki przeżycia; osobniki chcąc przeżyc muszą się do tych warunków przystosować,
 - 4. z 3 wynika, że najlepiej przystosowane osobniki mają największą szansę na przeżycie i rozmnażanie,
 - 5. w procesie rozmnażania osobniki najczęściej przekazują swoje "przystosowanie" potomstwu nie jest to regułą i możliwe jest wydanie potomstwa gorzej niż rodzice przystosowanego do warunków określonych przez środowisko,
 - 6. oprócz cech poprawiających przystosowanie, u potomstwa mogą występować losowe zmiany (*mutacje*), które mogą prowadzić do zmian ich przystosowania (lepsze lub gorsze przystosowanie),
 - 7. powyższe (jeśli środowisko nie zmieni diamatralnie warunków przystosowania) prowadzi (w czasie) do coraz lepszego przystosowania populacji.

- 1. Wprowadzenie do AG
- b) podstawowe definicje
 - populacja: zbiór osobników określonej liczebności.
 - osobnik: zakodowany w postaci chromosomów zbiór parametrów zadania rozwiązanie, punkty przeszukiwanej przestrzeni.
 - chromosom(y): ciągi kodowe składające się z genów.
 - gen: cecha, pojedynczy element genotypu.
 - genotyp: zespół chromosomów osobnika (struktra osobnika). Osobniki mogą być genotypami lub chromosomami, jeśli gentotyp składa się tylko z jednego chromosomu.
 - fenotyp: zestaw wartości odpowiadających danemu genotypowi.
 - allel: wartość danego genu (wartość cechy).
 - locus: pozycja wskazująca miejsce położenia danego genu w ciągu (chromosomie).

- 1. Wprowadznie do AG
- c) operatory genetyczne

selekcja - polega na wyborze z bieżącej populacji osobników, których materiał genetyczny zostanie poddany operacji krzyżowania oraz mutacji i przekazny osobnikom potomnym (kolejna populacja). Wybór następuje na podstawie określonej metody selekcji. Osobniki oceniane są na podstawie tzw. funkcji przystosowania/dopasowania (ang. fitness function - ff). Selekcja powinna promować osobniki najlepiej przystosowane - o najwyższej wartości ff

krzyżowanie (ang. crossover) - polega na wymianie materiału genetycznego pomiędzy losowo wybranymi (podczas selekcji) parami osobników. W wyniku krzyżowania powstają nowe osobniki, które mogą wejść w skład nowej populacji (kolejnego pokolenia). Osobniki powstałe po krzyżowaniu powinny być - i są, jeśli metoda selekcji jest właściwa - lepiej przystosowane (wyższa wartość ff) od swoich rodziców. Zachodzi z prawdopodobieństwem p_c .

 $\mathbf{mutacja}$ - polega na zamianie wartości losowo wybranego genu (cechy osobnika). Celem użycia operatora mutacji jest zapewnienie zmienności chromosomów. W przypadku wykorzystania AG do poszukiwania rozwiązania np. problemów kombinatorycznych stwarza możliwość wyjścia z optimów lokalnych ilub zwiększenia intensyfikacji przeszukiwania. Zachodzi z prawdopodobieństwem p_m .

Pytanie 1. Jakie są efekty użycia operatorów kzyżowania i mutacji?

- 1. Wprowadzenie do AG
- d) konstruowanie AG

Do skonstruowania AG dla dowolnego zadania wymagane jest sześć składowych:

- 1. genetyczna reprezentacja potencjalnych rozwiązań (osobników).
- 2. funkcja oceniająca jakość rozwiązania (funkcja przystosowania).
- 3. metoda generowania populacji początkowej.
- 4. operatory genetyczne (metoda selekcji, opertarory krzyżowania i mutacji).
- 5. parametrów charakteryzujacych algorytm: $rozmiar\ populacji,\ p_c,\ p_m.$
- 6. warunków zakończenia działania, tzw. warunków stopu/zatrzymania.

Wprowadzenie do AG

- e) standardowy AG
- 1. wybór populacji początkowej chromosomów (losowy)
- 2. ocena przystosowania chromosomów
- 3. sprawdzanie warunku zatrzymania
 - a. selekcja chromosomów wybór populacji macierzystej (ang. mating pool)
 - b. krzyżowanie chromosomów z populacji rodzicielskiej
 - c. mutacja może być również wykonana przed krzyżowaniem
 - d. ocena przystosowania chromosomów
 - e. utworzenie nowej populacji
- 4. wyprowadzenie ,,najlepszego'' rozwiązania

Pytanie 2. Jaki przebiega typowy algorytm optymalizujący funkcję jednej zmiennej?

Pytanie 3. Jak wygląda wykres wartości dopasowania w czasie?

Wprowadzenie do AG

f) przykład rozwiązania - dyskretny problem plecakowy (1)

i	$ig p_i$	w_i	p_i/w_i
1	40	2	20
2	30	5	6
3	50	10	5
4	10	5	2

$$n = 4, W = 16$$

Wprowadzenie do AG

f) przykład rozwiązania - dyskretny problem plecakowy (2)

- postać rozwiązania: $x = [x_1, x_2, x_3, x_4], x \in 0, 1$
- optymalizowana funkcja: $f(x) = \sum_{i=1}^{4} x_i \cdot w_i$
- potencjalne rozwązania: $x^1 = [1, 0, 0, 0]; x^2 = [1, 0, 0, 1]; x^3 = [0, 0, 0, 0]; x^4 = [1, 0, 1, 0]; x^5 = [1, 1, 1, 0]$
- rozwązania dopuszczalne: $x^1 = [1,0,0,0]; x^2 = [1,0,0,1]; x^3 = [0,0,0,0]; x^4 = [1,0,1,0]$
- wartości $ff: ff(x^1) = 40; ff(x^2) = 50; ff(x^3) = 0; ff(x^4) = 90$
- średnia wartość ff dla populacji w pokoleniu 1: ff = 45
- $\bullet \;\;$ populacja macierzysta: $x^1 = [1,0,0,0]; x^2 = [1,0,0,1]; x^4 = [1,0,1,0]$
- pary do reprodukcji: $C_1(x^1 = [1, 0, 0, 0] : x^4 = [1, 0, 1, 0]); C_2(x^2 = [1, 0, 0, 1] : x^4 = [1, 0, 1, 0])$
- krzyżowanie $loci: C_1: 2, C_2: 3 (p_c = 1): \dots$

• ...

- pary do reprodukcji: $C_1(x^1 = [1, 0, 0, 0] : x^4 = [1, 0, 1, 0]); C_2(x^2 = [1, 0, 0, 1] : x^4 = [1, 0, 1, 0])$
- $\bullet \quad \text{krzyżowanie} \ \textit{loci:} \ C_1:2, C_2:3 \ (p_c=1): \ x^{14}=[1,0|1,0]; x^{41}=[1,0|0,0]; x^{24}=[1,0,0|0]; x^{42}=[1,0,1|1]$
- wartości $ff: ff(x^{14}) = 90; ff(x^{41}) = 40; ff(x^{24}) = 40; ff(x^{42}) = 80$
- nowe pokolenie: $x^4 = [1, 0, 1, 0]; x^1 = [1, 0, 1, 0]; x^2 = [1, 0, 0, 1]; x^{42} = [1, 0, 1, 1]$
- średnia wartość ff dla populacji w pokoleniu 2: $ff = 62.5 \longrightarrow ff(2) > ff(1)$
- powtarzane, aż do wystąpienia warunku zatrzymania.

Wprowadzenie do AG

g) nieco historii

- opracowane w latach 70-tych w USA.
- twórcą AG jest John H. Holland (1929 2015) Adaptation in Natural and Artificial Systems, 1975.
- David E. Goldberg Genetic Algorithms in Search, Optimazation and Machine Learning, 1989

Wprowadzenie do AG

f) literatura

- David E. Goldberg Algorytmy genetyczne i ich zastosowania, 1998
- Z. Michalewicz Algorytmy genetyczne + struktury danych = programy ewolucyjne, 1996
- John R. Koza Genetic Programming I, II, III, IV . . . , 1992, 1994, 1999, 2003
- J. Arabas Wykłady z algorytmów ewolucyjnych, 2001

2. Algorytmy ewolucyjne (AE) (1)

Podział AE wg sposobu kodowania rozwiązań:

- algorytmy genetyczne osobnik w postaci ciągu bitów o stałej długości.
- strategie ewolucyjne osobnik jest wektorem liczb rzeczywistych $x = [x_1, \dots, x_n], x \in \mathbb{R}^n$.
- programowanie genetyczne osobnik jest drzewem reprezentującym program komputerowy.
- programowanie ewolucyjne osobnik jest wektorem liczb rzeczywistych brak krzyżowania; tylko mutacja.

2. Algorytmy ewolucyjne (AE) (2)

Charakterystyczne dla AE są:

- użycie populacji (algorytmy populacyjne),
- stochastyczne operatory przeszukiwania,
- reprodukcja odbywa się na podstawie wartości funkcji przystosowania.

Osobnik: chromosom złożony z dwóch genów (równej długości)

 $\mathtt{genotyp} \colon 0101010010101 - 0101010010101$

 \downarrow dekodowanie

fenotyp x: 0.56 - 1.45 - $x_1|x_2$

 \downarrow obliczenie dopasowania

$$ff(x_1, x_2) = 3.745$$

Funkcja dopasowania (f(x))

- pełni rolę środowiska, ocenia potencjalne rozwiązania.
- jest odwzorowaniem ciągów binarnych o długości B w zbiór \mathbb{R}^+ .
- może być użyta bezspośrednio jako funkcja celu (problemy optymalizacyjne), jeśli funkcja celu przyjmuje wartości dodatnie i problem jest problemem maksymalizacji.
- ullet jeśli problem jest problemem minimalizacji i $funkcja\ celu\ (g(x))$ może przyjmować wartości ujemne, wtedy

$$f(x) = g(x) + C_1$$

ullet jeśli problem jest problemem maksymalizacji i $funkcja\ celu\ (g(x))$ może przyjmować wartości ujemne, wtedy

$$f(x) = C_2 - g(x)$$

Selekcja

- Zgodnie z ewolucją darwinowską, lepiej przystosowani mają (powinni mieć) większe szanse na rozmnażanie.
- W związku z tym stosujemy selekcję proporcjonalną. Losujemy liczbę, z przedziału [0,100] i z prawdopodobieństwem

$$p(x) = \frac{f(x)}{\sum_{i=1}^{S} f(x_i)}$$

Krzyżowanie

- Dwuargumentowy operator krzyżowania stosowany jest dla par osobników łączenie w pary jest losowe z puli rodzicielskiej (wynik selekcji).
- Krzyżowanie realizowane jest w prawdopodobieństwem p_c . Typowo 0.6 do 1.0, ale . . . zależy od problemu oraz możliwego wpływu innych parametrów, np.: p_m .
 - wybierz losowo punkt krzyżowania (locus),
 - podziel genotypy rodziców wg. punktu krzyżowania,
 - połącz fragmenty "na krzyż" tworząc nowe osobniki.

Mutacja

- $\bullet\,$ Dla każdego osobnika populacji po krzyżowaniu dla każdego genu, z prawdopodobieństwem $p_m,$ zmień wartość na przeciwną.
- Mutacja realizowana jest z p_m . Typowo 0.01 do 0.1. Mutacja powinna prowadzić do niewielkich "kosmetycznych" acz znaczących zmian w genotypie osobnika.

Problemy

- krzyżowanie i mutacja działają tylko dla ciągów binarnych,
- model sukcesji (wszyscy rodzice umierają, potomkowie zastępują rodziców) sprawia, że najlepsze dopasowanie nie musi monotonicznie rosnąć w kolejnych pokoleniach,
- selekcja proporcjonalna może powodować, że we wczesnych fazach algorytmu pojawiają się superosobnicy.
- pojawia się zjawisko przedwczesnej zbieżności (ang. premature convergence).
- możliwe pojawienie się identycznego dopasowania (często dość szybko) całej populacji na poziomie średnim lub poziomie superosobnika (silna dominacja).
- pojawia się tzw. "błądzenie przypadkowe wśród przeciętniaków" (D.E. Goldberg).

Jak poprawić? (1)

- kryterium zatrzymania
 - po zadanej liczbie pokoleń **problemem jest tu ...**,
 - po osiągnięciu zadanej wartości funkcji dopasowania **problemem jest tu ...**,
 - gdy najlepsza wartość dopasowania nie ulega poprawie **problemem jest tu ...**,
 - po zakończeniu ustalonego czasu pracy **problemem jest tu ...**,

Jak poprawić? (2)

- selekcja turniejowa,
- selekcja rangowa,
- selekcja rangowa z liniową funkcja prawdopodobieństwa.

Selekcja turniejowa

Turniej:

- $\bullet\,\,$ wybierz losowo k~(k < L,gdzie L liczność populacji P)osobników (bez zwracania) z populacji P.
- ullet zachowaj najlepszego umieść w populacji S.
- zwróć k osbników do P^i .
- wykonaj L razy, aby |S| = L.

4. SAG a TSP

Reprezentacja rozwiązań

- 1. reprezentacja ścieżkowa (ang. path representation) kodowanie permutacyjne
- 2. reprezentacja porządkowa (ang. ordinal representation)

Reprezentacja ścieżkowa:

- w osobniku zapisane są kolejne miasta na trasie,
- wymaga specjalnych (niestandardowych) operatorów krzyżowania i mutacji.

Reprezentcja porządkowa:

- w osobniku zapisana jest kolejność wyboru miast (wierzchołków) z listy początkowej,
- ... czyli czai się pytanie ©

4. SAG a TSP

Reprezentacja ścieżkowa - operatory krzyżowania

- PMX Partially Mapped Crossover (Goldberg, 1985),
- OX Ordered Crossover (Davis, 1985),
- EX Edge Crossover (Whitley, 1989),
- SXX Subtour Exchanged Crossover (Yamamura, 1992),
- PX Partition Crossover (Whitley, 2000).

PMX - krzyżowanie z częściowym odwzorowaniem; odmiana krzyżowania dwupunktowego

Osobniki rodzicielskie

$$p = (p_1 \dots p_{k_1} \dots p_{k_2}) \dots p_m)$$
 $q = (q_1 \dots q_{k_1} \dots q_{k_2}) \dots q_m)$
 $p = (q_1 \dots q_{k_2} \dots q_{k_2}) \dots q_m)$

Osobniki potomne

$$sekcja \ dopasowania$$
 $r = (r_1 \dots \overbrace{|p_{k_1} \dots p_{k_2}|} \quad \dots \quad r_m)$
 $s = (s_1 \dots \underbrace{|q_{k_1} \dots q_{k_2}|} \quad \dots \quad s_m)$
 $sekcja \ dopasowania$

Indeksy k_1 i k_2 wybrane losowo z [1...m]. Wykonujemy transpozycje w obrębie chromosomu.

PMX - przykład

Osobniki rodzicielskie

$$p = (1 \ 2 \ | \ 3 \ 4 \ 5 \ 6 \ | \ 7 \ 8 \ 9)$$

$$q = (5 \ 3 \ | \ 6 \ 7 \ 8 \ 1 \ | \ 2 \ 9 \ 4)$$

Sekcja dopasowania: $k_1 = 3, k_2 = 7$

Tabela odwzorowań: $[3 \leftrightarrow 6], [4 \leftrightarrow 7], [5 \leftrightarrow 8], [6 \leftrightarrow 1]$

Osobniki potomne

$$p = (1 2 | 3 4 5 6 | 7 8 9)$$

 $q = (5 3 | 6 7 8 1 | 2 9 4)$

Wstawiamy miasta nie powodujące konfliktów (z osobników p i q)

Kolejne miasta wstawiamy zgodnie z tabelą odwzorowań $[3\leftrightarrow 6], [4\leftrightarrow 7], [5\leftrightarrow 8], [6\leftrightarrow 1]$. Otrzymujemy:

$$s = (8 - | 3 4 5 6 | 2 9 7)$$

Uzupełnienie pozostałych pozycji wymaga dwukrotnego zastosowania odwzorowań: $[6 \leftrightarrow 1]$ i $[3 \leftrightarrow 6]$.

Z czego otrzymujemy osobniki potomne:

$$r = (3 \ 2 \ | 6 \ 7 \ 8 \ 1 \ | 4 \ 5 \ 9)$$

$$s = (8 1 | 3 4 5 6 | 2 9 7)$$

$$p = (1 \ 2 \ | \ 3 \ 4 \ 5 \ 6 \ | \ 7 \ 8 \ 9)$$

$$q = (5 \ 3 \ | \ 6 \ 7 \ 8 \ 1 \ | \ 2 \ 9 \ 4)$$

OX - krzyżowanie z częściowym odwzorowaniem; odmiana krzyżowania dwupunktowego

Osobniki rodzicielskie

$$p = (p_1 \dots p_{k_1} \dots p_{k_2}) \dots p_m)$$
 $q = (q_1 \dots q_{k_1} \dots q_{k_2}) \dots q_m)$
 $p = (q_1 \dots q_{k_2} \dots q_{k_2}) \dots q_m)$

Osobniki potomne

$$sekcja\ dopasowania$$
 $r = (r_1 \dots \overbrace{ [p_{k_1} \dots p_{k_2}] } \dots r_m)$
 $s = (s_1 \dots \underbrace{ [q_{k_1} \dots q_{k_2}] }_{sekcja\ dopasowania} \dots s_m)$

Indeksy k_1 i k_2 wybrane losowo z [1...m]. Kopiujemy do r te elementy z x, które nie występują w sekcji dopasowania skopiowanej z q - począwszy od pozycji $k_2 + 1$.

OX - przykład

Osobniki rodzicielskie

$$p = (1 \ 2 \ | \ 3 \ 4 \ 5 \ 6 \ | \ 7 \ 8 \ 9)$$

$$q = (5 \ 3 \ | \ 6 \ 7 \ 8 \ 1 \ | \ 2 \ 9 \ 4)$$

Osobniki potomne

$$r = (\ 4 \ 5 \ | \ 6 \ 7 \ 8 \ 1 \ | \ 9 \ 2 \ 3 \)$$

$$s = (8 1 | 3 4 5 6 | 2 9 7)$$

EX - algorytm

- 1. Utwórz listy sąsiedztwa dla wszystkich wierzchołków; (+) osiągalny z obu rodziców
- 2. v przypisz losowy wierzchołek z dowolnego rodzica Dopóki NOWY (osobnik) nie jest utworzony, powtarzaj:
 - Dołącz v do NOWY
 - Usuń v z list sąsiedztwa
 - Jeżeli lista wierzchołka v jest pusta
 - = przypisz u losowo wierzchołek nie będący w NOWY
 - W przeciwnym przypadku
 - = wybierz z listy wierzchołek oznaczony (+) lub ten o najkrótszej liście
 - = jeśli wierzchołki z listy v mają równie długie listy
 - ≡ wybierz jeden losowo
 - = u przypisz wybrany wierzchołek

EX - przykład (1)

$$p_1 = (1 2 3 4 5 6 7 8 9)$$

$$p_2 = (9 \ 3 \ 7 \ 8 \ 2 \ 6 \ 5 \ 1 \ 4)$$

wierzchołek	krawędzie	wierzchołek	krawędzie
1	2,5,4,9	6	2,5,+,7
2	1,3,6,8	7	3,6,8+
3	2,4,7,9	8	2,7+,9
4	1,3,5,9	9	1,3,4,8
5	1,4,6+		

EX - przykład (2)

wybór	krawędź	powód wyboru	NOWY
all	1	losowy	1
2,5,4,9	5	najkrótsza lista	1 5
4,6	6	wspólna krawędź (+)	1 5 6
2,7	2	losowo (równa długość list)	1 5 6 2
3,8	8	najkrótsza lista	1 5 6 2 8
7,9	7	wspólna krawędź (+)	156287
3	3	jedna krawędź	1562873
4,9	9	losowo	15628739
4	4	ostatnia krawędź	156287394

4. SAG a TSP

Reprezentacja ścieżkowa - operatory mutacji

• inversion - wybiera losowo podciągi miast i zamienia ich kolejność.

$$p = (1 \ 2 \ | \ 3 \ 4 \ 5 \ 6 \ | \ 7 \ 8 \ 9 \) \rightarrow p' = (1 \ 2 \ | \ 6 \ 5 \ 4 \ 3 \ | \ 7 \ 8 \ 9 \)$$

• insertion - przestawia losowo wybrane miasto na inną (losowo wybraną) pozycję, "rozsuwając" pozostałe.

$$q = (1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8 \ 9) \rightarrow q' = (1 \ 2 \ 7 \ 3 \ 4 \ 5 \ 6 \ 8 \ 9)$$

• displacement - zamienia w wybranym losowowo podciągu miast pierwsze z ostatnim

$$r = (1 \ 2 \ | \ 3 \ 4 \ 5 \ 6 \ | \ 7 \ 8 \ 9 \) \rightarrow r' = (1 \ 2 \ | \ 6 \ 4 \ 5 \ 3 \ | \ 7 \ 8 \ 9 \)$$

• transposition - zamienia dwa losowo wybrane miasta

$$s = (1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8 \ 9) \rightarrow s' = (7 \ 2 \ 3 \ 4 \ 5 \ 6 \ 1 \ 8 \ 9)$$

4. SAG a TSP

Reprezentacja porządkowa

Osobnik reprezentuje kolejność wyboru z pewnej listy początkowej miast tworzących trasę. Kolejne geny określają, które z **z postałych** miast należy wybrać jako następne.

lista miast:

$$[1 2 3 4 5 6 7 8 9]$$

osobnik:

$$p = (1 1 2 1 4 1 3 1 1)$$

trasa:

$$<$$
 1 2 4 3 8 5 9 6 7 $>$

Pytanie. Jakie operatory genetyczne są wymagane do uzyskiwania osobników dopuszczalnych?