

Algorytmy i modele inspirowane biologicznie

2. Teoria algorytmów ewolucyjnych

Maciej Komosiński



**Fundusze
Europejskie**
Polska Cyfrowa



**Rzeczpospolita
Polska**

Unia Europejska
Europejski Fundusz
Rozwoju Regionalnego



Przypomnienie z poprzedniego semestru (i ew. uzupełnienie)

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

- O czym mówi *Twierdzenie o schematach* Hollanda?
- O czym mówi *Hipoteza cegiełek*?
- O czym mówi *Twierdzenie „No Free Lunch”*?

Budowa problemów zawodnych (zwodniczych)

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

Aby skonstruować taki problem, który sprawi trudność algorytmowi genetycznemu, zanegujmy hipotezę cegiełek: dobre „cegiełki” po połączeniu mają tworzyć niekorzystną strukturę.

Najprostszy taki przypadek można uzyskać dla genotypów o długości 2 (jest to problem zwodniczy rzędu drugiego).

Budujemy problem zwodniczy rzędu drugiego

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

Założmy, że istnieją cztery schematy:

```
*** 0 *** 0 ***  
*** 0 *** 1 ***  
*** 1 *** 0 ***  
*** 1 *** 1 ***
```

Gwiazdki odpowiadają dowolnej liczbie symboli nieustalonych (ale pozycje ustalone są we wszystkich schematach na tych samych miejscach). Oznaczmy średnie oceny schematów przez f_{00} , f_{01} , f_{10} , f_{11} i schemat z dwoma jedynekami jest globalnym optimum (f_{11}). Aby problem był zwodniczy, chcemy, by schematy z jednym ustalonym zerem były lepsze od odpowiednich schematów z jedyneką – to znaczy, by spełniona była przynajmniej jedna z nierówności:

$$\begin{aligned} f_{0*} &> f_{1*} & f_{0*} &= (f_{00} + f_{01})/2, \text{ itd.} \\ f_{*0} &> f_{*1} \end{aligned}$$

Krajobraz przystosowania minimalnego problemu zwodniczego

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

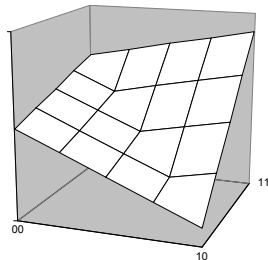
Bibliografia

Taki problem nazywany jest minimalnym problemem zwodniczym (ang. *minimal deceptive problem*, *MDP*), ponieważ nie istnieje problem zwodniczy rzędu pierwszego.

Są dwa typy problemów zwodniczych rzędu drugiego (wykres obok pokazuje przykładowy problem zwodniczy rzędu drugiego, typu I):

Typ I: $f_{01} > f_{00}$

Typ II: $f_{00} \geq f_{01}$



W *MDP* funkcja przystosowania *nie może* być przedstawiona jako liniowa kombinacja poszczególnych alleli, czyli w postaci
$$f(x_1, x_2) = b + \sum_{i=1}^2 a_i x_i.$$

Trudność problemów zwodniczych w praktyce

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

To, że problem jest zwodniczy nie oznacza jeszcze, iż algorytm genetyczny nie znajdzie optimum. Oznacza to jednak, że funkcja przystosowania (przedstawiona na osi pionowej) nie może być wyrażona jako liniowa kombinacja poszczególnych bitów, zatem występuje zjawisko epistazy – nieliniowości. Zwykle taki problem nie okazuje się jednak AG-trudny, czyli algorytm genetyczny może znaleźć rozwiązanie optymalne.

Zachowanie algorytmu zależy jednak od wielu czynników, takich jak na przykład początkowa obecność schematów w populacji osobników. W szczególności, jeśli wszystkie cztery schematy występują w populacji początkowej, problem zwodniczy rzędu 2 typu I nie jest AG-trudny. Dla zadań typu II zachowanie algorytmu zależy od proporcji występowania schematów w populacji: jeśli schemat 00 będzie występował w przewadze, algorytm może być zbieżny do rozwiązania nieoptymalnego (choć taka sytuacja występuje rzadko). Bardziej szczegółowa analiza problemów zwodniczych znajduje się w [Gol03, str. 63–70, 375–382].

Cecha reprezentacji (i ew. operatorów dla niej) – stopień zależności pomiędzy różnymi genami w chromosomie [Dav90]. Jeśli dana reprezentacja posiada wysoką epistazę, fenotypowy efekt pewnych genów zależy od alleli (wartości) innych genów (poligeniczność).

Zerowa epistaza: każdy gen niezależnie wpływa na wartość funkcji celu (i wtedy nie warto stosować AE). Dla skuteczności AE **tym lepiej**, im mniejsza epistaza. Dla niektórych definicji funkcji celu, projektując reprezentację i operatory może opłacać się zaakceptować niewielki wzrost epistazy, jeśli zyskamy korzystniejszy związek między topologią przestrzeni przeszukiwania a krajobrazem przystosowania.

Oszacuj mniej więcej epistazę dla następujących funkcji celu $f(x_1, x_2)$, gdzie x to liczby – wartości genów: $x_1 + x_2$, $x_1 - x_2$, $x_1 \cdot x_2$, $x_1 + x_2 \cdot x_2$, $x_1 + x_1 \cdot x_2$, $\frac{x_1}{x_2}$, $x_2 + \frac{x_1}{x_2}$.

Epistaza – przykład

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

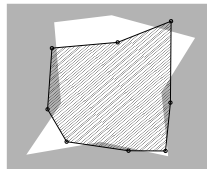
Brak mutacji

Zadanie

Bibliografia

Optymalizujemy kształt figury 2D (albo przekrój obiektu 3D); geny opisują rozmieszczenie wierzchołków w przestrzeni (współrzędne x_i, y_i).

Porównaj epistazę dla następujących scenariuszy:



- 1 Same geny x_i, y_i oraz operator mutacji przesuujący losowy wierzchołek.
- 2 Wprowadzamy dwa dodatkowe geny: obrotu i skali. Mutacja tych pojedynczych genów zmienia orientację (obróć) i rozmiar całej figury.
- 3 Zamiast tych dwóch dodatkowych genów wprowadzamy specjalne operatory mutacji-obrotu i mutacji-skalowania.

Epistaza – przykład

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

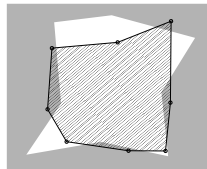
Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

Optymalizujemy kształt figury 2D (albo przekrój obiektu 3D); geny opisują rozmieszczenie wierzchołków w przestrzeni (współrzędne x_i, y_i).



Porównaj epistazę dla następujących scenariuszy:

- 1 Same geny x_i, y_i oraz operator mutacji przesuający losowy wierzchołek.
- 2 Wprowadzamy dwa dodatkowe geny: obrotu i skali. Mutacja tych pojedynczych genów zmienia orientację (obrót) i rozmiar całej figury.
- 3 Zamiast tych dwóch dodatkowych genów wprowadzamy specjalne operatory mutacji-obrotu i mutacji-skalowania.

W scenariuszu (1) zmiana orientacji i rozmiaru figury jest możliwa jedynie przez bardzo wiele niezależnych mutacji współrzędnych wierzchołków. Tymczasem w pewnych zadaniach (np. w optymalizacji online, gdy docelowy kształt otworu się zmienia), możliwość szybkiego obrotu i skalowania może przyspieszyć zbieżność i polepszyć jakość uzyskiwanych rozwiązań. W scenariuszu (2) rozważ wpływ konkretnej implementacji (matematyczne działania przekształcające genotyp $\rightarrow$ fenotyp) – chodzi o efekty złożenia mutacji (1) i (2).

Epistaza – opis formalny

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

Epistaza jako stopień interakcji pomiędzy genami [RW95; NK00]:

$$f(s) = \text{stała}$$

$$+ \sum_{i=0}^{l-1} \text{efekt } s_i$$

$$+ \sum_{i=0}^{l-2} \sum_{j=i+1}^{l-1} \text{interakcja między } s_i \text{ i } s_j$$

$$+ \dots$$

$$+ \text{interakcja między } s_0, s_1, \dots, s_{l-1}$$

$$+ \text{losowy błąd.}$$

s_i – gen i -ty (lub **allel** genu i -tego, jeśli s jest konkretnym rozwiązaniem),
 l – liczba genów (numerowane od zera).

Interakcje mogą być dodatnie i ujemne: jeśli np. interakcja pomiędzy s_i i s_j ma taki sam znak, co efekty s_i i s_j , to taka interakcja wzmacnia działanie tych dwóch genów i może być pożądana.

Epistaza i mutacje

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

Zmienianie (mutacja) wartości genów powiązanych epistatycznie wpływa na jakość rozwiązania zależnie od rodzaju epistazy.

Uśrednione konsekwencje ujemnej, zerowej i dodatniej epistazy ilustruje [ten wykres](#).

Empiryczna i teoretyczna ocena AE

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

Badanie efektywności i zachowania AE można prowadzić teoretycznie albo eksperymentalnie (**empirycznie**).

Analiza teoretyczna daje ugruntowaną, pewną wiedzę, ale często da się tak badać tylko bardzo proste modele (tj. z wieloma założeniami).

Analiza empiryczna AE prowadzi do mniej pewnej wiedzy, trudniej o uogólnienia, ale zawsze da się taką analizę przeprowadzić. Oprócz testowania skuteczności algorytmu na docelowym problemie optymalizacji, często używa się znanych problemów testowych* albo sparametryzowanych modeli problemów pozwalających na sterowanie nasileniem epistazy – takich jak model NK** oraz jego warianty z neutralnością: NKP (probabilistyczny) i NKQ (z kwantyzacją).

*https://en.wikipedia.org/wiki/Test_functions_for_optimization

**https://en.wikipedia.org/wiki/NK_model

Konsekwencje neutralności w krajobrazach przystosowania [Gea+02]

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

- potocznie/tradycyjnie uważa się, że krajobraz przystosowania „składa się z pagórków”, a mutacje powodują zmianę przystosowania

Konsekwencje neutralności w krajobrazach przystosowania [Gea+02]

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

- potocznie/tradycyjnie uważa się, że krajobraz przystosowania „składa się z pagórków”, a mutacje powodują zmianę przystosowania
- prowadzi to do koncepcji „optimów lokalnych”, w których rozwiązania lub ich populacje mogą utknąć

Wstęp

Problemy zawodne

Epistaza

Ocena AE

Neutralność

Dryf genetyczny

Brak mutacji

Zadanie

Bibliografia

- potocznie/tradycyjnie uważa się, że krajobraz przystosowania „składa się z pagórków”, a mutacje powodują zmianę przystosowania
- prowadzi to do koncepcji „optimów lokalnych”, w których rozwiązania lub ich populacje mogą utknąć
- badania molekularne nad krajobrazami różnych sposobów zwijania RNA sugerują, że duża część mutacji na poziomie molekularnym jest selekcyjnie neutralna

Konsekwencje neutralności w krajobrazach przystosowania [Gea+02]

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

- potocznie/tradycyjnie uważa się, że krajobraz przystosowania „składa się z pagórków”, a mutacje powodują zmianę przystosowania
- prowadzi to do koncepcji „optimów lokalnych”, w których rozwiązania lub ich populacje mogą utknąć
- badania molekularne nad krajobrazami różnych sposobów zwijania RNA sugerują, że duża część mutacji na poziomie molekularnym jest selekcyjnie neutralna
- neutralność jest również obecna w wielu rzeczywistych problemach optymalizacyjnych (popularne „płaskowyzę” lub sąsiedzi tej samej jakości)

Konsekwencje neutralności w krajobrazach przystosowania [Gea+02]

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

- potocznie/tradycyjnie uważa się, że krajobraz przystosowania „składa się z pagórków”, a mutacje powodują zmianę przystosowania
- prowadzi to do koncepcji „optimów lokalnych”, w których rozwiązania lub ich populacje mogą utknąć
- badania molekularne nad krajobrazami różnych sposobów zwijania RNA sugerują, że duża część mutacji na poziomie molekularnym jest selekcyjnie neutralna
- neutralność jest również obecna w wielu rzeczywistych problemach optymalizacyjnych (popularne „płaskowyzę” lub sąsiedzi tej samej jakości)
- zatem: wiele genotypów [→ jeden fenotyp] → jedna wartość funkcji celu

Konsekwencje neutralności w krajobrazach przystosowania [Gea+02]

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

- potocznie/tradycyjnie uważa się, że krajobraz przystosowania „składa się z pagórków”, a mutacje powodują zmianę przystosowania
- prowadzi to do koncepcji „optimów lokalnych”, w których rozwiązania lub ich populacje mogą utknąć
- badania molekularne nad krajobrazami różnych sposobów zwijania RNA sugerują, że duża część mutacji na poziomie molekularnym jest selekcyjnie neutralna
- neutralność jest również obecna w wielu rzeczywistych problemach optymalizacyjnych (popularne „płaskowyzę” lub sąsiedzi tej samej jakości)
- zatem: wiele genotypów [→ jeden fenotyp] → jedna wartość funkcji celu
- neutralność: jedna z przyczyn równowag przestankowych

Konsekwencje neutralności w krajobrazach przystosowania [Gea+02]

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

- potocznie/tradycyjnie uważa się, że krajobraz przystosowania „składa się z pagórków”, a mutacje powodują zmianę przystosowania
- prowadzi to do koncepcji „optimów lokalnych”, w których rozwiązania lub ich populacje mogą utknąć
- badania molekularne nad krajobrazami różnych sposobów zwijania RNA sugerują, że duża część mutacji na poziomie molekularnym jest selekcyjnie neutralna
- neutralność jest również obecna w wielu rzeczywistych problemach optymalizacyjnych (popularne „płaskowyzę” lub sąsiedzi tej samej jakości)
- zatem: wiele genotypów [→ jeden fenotyp] → jedna wartość funkcji celu
- neutralność: jedna z przyczyn równowag przestankowych
- jeśli neutralność występuje często w krajobrazie, ryzyko uwięzienia populacji AE w lokalnych optimach jest niskie

Konsekwencje neutralności w krajobrazach przystosowania [Gea+02]

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

- potocznie/tradycyjnie uważa się, że krajobraz przystosowania „składa się z pagórków”, a mutacje powodują zmianę przystosowania
- prowadzi to do koncepcji „optimów lokalnych”, w których rozwiązania lub ich populacje mogą utknąć
- badania molekularne nad krajobrazami różnych sposobów zwijania RNA sugerują, że duża część mutacji na poziomie molekularnym jest selekcyjnie neutralna
- neutralność jest również obecna w wielu rzeczywistych problemach optymalizacyjnych (popularne „płaskowyzę” lub sąsiedzi tej samej jakości)
- zatem: wiele genotypów [→ jeden fenotyp] → jedna wartość funkcji celu
- neutralność: jedna z przyczyn równowag przestankowych
- jeśli neutralność występuje często w krajobrazie, ryzyko uwięzienia populacji AE w lokalnych optimach jest niskie
- w związku z tym rośnie rola charakterystyk operatorów rekonfiguracji (mutacji/sąsiedztwa, krzyżowania) i rola dryfu genetycznego

Wstęp

Problemy zawodne

Epistaza

Ocena AE

Neutralność

Dryf genetyczny

Brak mutacji

Zadanie

Bibliografia

- potocznie/tradycyjnie uważa się, że krajobraz przystosowania „składa się z pagórków”, a mutacje powodują zmianę przystosowania
- prowadzi to do koncepcji „optimów lokalnych”, w których rozwiązania lub ich populacje mogą utknąć
- badania molekularne nad krajobrazami różnych sposobów zwijania RNA sugerują, że duża część mutacji na poziomie molekularnym jest selekcyjnie neutralna
- neutralność jest również obecna w wielu rzeczywistych problemach optymalizacyjnych (popularne „płaskowyzę” lub sąsiedzi tej samej jakości)
- zatem: wiele genotypów [$\rightarrow$ jeden fenotyp] $\rightarrow$ jedna wartość funkcji celu
- neutralność: jedna z przyczyn równowag przestankowych
- jeśli neutralność występuje często w krajobrazie, ryzyko uwięzienia populacji AE w lokalnych optimach jest niskie
- w związku z tym rośnie rola charakterystyk operatorów rekonfiguracji (mutacji/sąsiedztwa, krzyżowania) i rola dryfu genetycznego
- kluczowa rola ostrej i nieostrej nierówności w implementacjach przeszukiwania lokalnego, o której mówiliśmy wcześniej (ostra nierówność $\rightarrow$ wszystkie neutralne ruchy są potencjalnymi końcami)

Główne siły prowadzące ewolucję

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

- co kieruje ewolucją? co wpływa na trajektorię populacji?

Główne siły prowadzące ewolucję

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

- co kieruje ewolucją? co wpływa na trajektorię populacji?
- krajobraz przystosowania kontra operatory rekonfiguracyjne

Główne siły prowadzące ewolucję

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

- co kieruje ewolucją? co wpływa na trajektorię populacji?
- krajobraz przystosowania kontra operatory rekonfiguracyjne
- co się stanie, gdy wszystkie lub większość mutacji będzie niekorzystna, a dawne osobniki nie będą zachowywane?

Główne siły prowadzące ewolucję

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

- co kieruje ewolucją? co wpływa na trajektorię populacji?
- krajobraz przystosowania kontra operatory rekonfiguracyjne
- co się stanie, gdy wszystkie lub większość mutacji będzie niekorzystna, a dawne osobniki nie będą zachowywane?
- „bias” / „neutralność” krajobrazu przystosowania vs. bias/neutralność operatorów rekonfiguracyjnych

Główne siły prowadzące ewolucję

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

- co kieruje ewolucją? co wpływa na trajektorię populacji?
- krajobraz przystosowania kontra operatory rekonfiguracyjne
- co się stanie, gdy wszystkie lub większość mutacji będzie niekorzystna, a dawne osobniki nie będą zachowywane?
- „bias” / „neutralność” krajobrazu przystosowania vs. bias/neutralność operatorów rekonfiguracyjnych
- operatory rekonfiguracyjne: zmiany w częstości alleli, ale zwykle chcemy uniknąć „biasu” !

Główne siły prowadzące ewolucję

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

- co kieruje ewolucją? co wpływa na trajektorię populacji?
- krajobraz przystosowania kontra operatory rekonfiguracyjne
- co się stanie, gdy wszystkie lub większość mutacji będzie niekorzystna, a dawne osobniki nie będą zachowywane?
- „bias” / „neutralność” krajobrazu przystosowania vs. bias/neutralność operatorów rekonfiguracyjnych
- operatory rekonfiguracyjne: zmiany w częstości alleli, ale zwykle chcemy uniknąć „biasu” !
- widzisz populację osób o niebieskich, zielonych i brązowych oczach; po ilu pokoleniach wszyscy mają zielone oczy. Dlaczego?

Dryf genetyczny

Wstęp	
Problemy zawodne	
Epistaza	
Ocena AE	
Neutralność	
Dryf genetyczny	• wyłączmy presję selekcyjną, mutację i krzyżowanie
Brak mutacji	
Zadanie	
Bibliografia	

Dryf genetyczny

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

- wyłączmy presję selekcyjną, mutację i krzyżowanie
- przykład kamyków w słoiku

Dryf genetyczny

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

- wyłączmy presję selekcyjną, mutację i krzyżowanie
- przykład kamyków w słoiku
- dyskretny charakter populacji (składających się z odrębnych osobników), więc idealnie równy rozkład alleli nie jest możliwy (np. 1/64 wśród 5000 osobników)

Dryf genetyczny

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

- wyłączmy presję selekcyjną, mutację i krzyżowanie
- przykład kamyków w słoiku
- dyskretny charakter populacji (składających się z odrębnych osobników), więc idealnie równy rozkład alleli nie jest możliwy (np. 1/64 wśród 5000 osobników)
- wąskie gardła (populacja tymczasowo kurczy się do bardzo małych rozmiarów)

Dryf genetyczny

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

- wyłączmy presję selekcyjną, mutację i krzyżowanie
- przykład kamyków w słoiku
- dyskretny charakter populacji (składających się z odrębnych osobników), więc idealnie równy rozkład alleli nie jest możliwy (np. 1/64 wśród 5000 osobników)
- wąskie gardła (populacja tymczasowo kurczy się do bardzo małych rozmiarów)
 - efekt założyciela

Dryf genetyczny

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

- wyłączmy presję selekcyjną, mutację i krzyżowanie
- przykład kamyków w słoiku
- dyskretny charakter populacji (składających się z odrębnych osobników), więc idealnie równy rozkład alleli nie jest możliwy (np. 1/64 wśród 5000 osobników)
- wąskie gardła (populacja tymczasowo kurczy się do bardzo małych rozmiarów)
 - efekt założyciela
- wpływ rozkładu cech i wielkości populacji

Dryf genetyczny

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

- wyłączmy presję selekcyjną, mutację i krzyżowanie
- przykład kamyków w słoiku
- dyskretny charakter populacji (składających się z odrębnych osobników), więc idealnie równy rozkład alleli nie jest możliwy (np. 1/64 wśród 5000 osobników)
- wąskie gardła (populacja tymczasowo kurczy się do bardzo małych rozmiarów)
 - efekt założyciela
- wpływ rozkładu cech i wielkości populacji
 - prawdopodobieństwo, że dana cecha ostatecznie zdominuje całą populację, to jej częstość w populacji

Dryf genetyczny

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

- wyłączmy presję selekcyjną, mutację i krzyżowanie
- przykład kamyków w słoiku
- dyskretny charakter populacji (składających się z odrębnych osobników), więc idealnie równy rozkład alleli nie jest możliwy (np. 1/64 wśród 5000 osobników)
- wąskie gardła (populacja tymczasowo kurczy się do bardzo małych rozmiarów)
 - efekt założyciela
- wpływ rozkładu cech i wielkości populacji
 - prawdopodobieństwo, że dana cecha ostatecznie zdominuje całą populację, to jej częstość w populacji
 - oczekiwana liczba pokoleń do wystąpienia całkowitej dominacji jest proporcjonalna do wielkości populacji

Dryf genetyczny

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

- wyłączmy presję selekcyjną, mutację i krzyżowanie
- [przykład kamyków w słoiku](#)
- dyskretny charakter populacji (składających się z odrębnych osobników), więc idealnie równy rozkład alleli nie jest możliwy (np. 1/64 wśród 5000 osobników)
- [wąskie gardła](#) (populacja tymczasowo kurczy się do bardzo małych rozmiarów)
 - [efekt założyciela](#)
- wpływ rozkładu cech i wielkości populacji
 - prawdopodobieństwo, że dana cecha ostatecznie zdominuje całą populację, to jej częstość w populacji
 - oczekiwana liczba pokoleń do wystąpienia całkowitej dominacji jest [proporcjonalna do wielkości populacji](#)
- wpływ losowości (nieograniczonej) vs. „sprawiedliwość” – przypomnienie: nasza wcześniejsza szczegółowa dyskusja

Dryf genetyczny

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

- wyłączmy presję selekcyjną, mutację i krzyżowanie
- przykład kamyków w słoiku
- dyskretny charakter populacji (składających się z odrębnych osobników), więc idealnie równy rozkład alleli nie jest możliwy (np. 1/64 wśród 5000 osobników)
- wąskie gardła (populacja tymczasowo kurczy się do bardzo małych rozmiarów)
 - efekt założyciela
- wpływ rozkładu cech i wielkości populacji
 - prawdopodobieństwo, że dana cecha ostatecznie zdominuje całą populację, to jej częstość w populacji
 - oczekiwana liczba pokoleń do wystąpienia całkowitej dominacji jest proporcjonalna do wielkości populacji
- wpływ losowości (nieograniczonej) vs. „sprawiedliwość” – przypomnienie: nasza wcześniejsza szczegółowa dyskusja

Dryf genetyczny

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

- wyłączmy presję selekcyjną, mutację i krzyżowanie
- [przykład kamyków w słoiku](#)
- dyskretny charakter populacji (składających się z odrębnych osobników), więc idealnie równy rozkład alleli nie jest możliwy (np. 1/64 wśród 5000 osobników)
- [wąskie gardła](#) (populacja tymczasowo kurczy się do bardzo małych rozmiarów)
 - [efekt założyciela](#)
- wpływ rozkładu cech i wielkości populacji
 - prawdopodobieństwo, że dana cecha ostatecznie zdominuje całą populację, to jej częstość w populacji
 - oczekiwana liczba pokoleń do wystąpienia całkowitej dominacji jest [proporcjonalna do wielkości populacji](#)
- wpływ losowości (nieograniczonej) vs. „sprawiedliwość” – przypomnienie: nasza wcześniejsza szczegółowa dyskusja

-
- jeśli nie znasz tych zjawisk, ich konsekwencje mogą być sprzeczne z intuicją, błędnie interpretowane i mylnie przypisywane innym mechanizmom!
 - interakcja krajobrazu przystosowania (selekcja), operatorów rekonfiguracji i dryfu genetycznego

Przykład analizy teoretycznej AG

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

Wyobraźmy sobie algorytm genetyczny (reprezentacja binarna) z selekcją ruletkową i krzyżowaniem, bez mutacji.

Co się stanie po długim czasie działania algorytmu?

Przykład analizy teoretycznej AG

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

Wyobraźmy sobie algorytm genetyczny (reprezentacja binarna) z selekcją ruletkową i krzyżowaniem, bez mutacji.

Co się stanie po długim czasie działania algorytmu?

Jak to udowodnić?

Przykład analizy teoretycznej AG

Stany i przejścia

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

Rozważamy populację m osobników o długości k . Ile jest różnych stanów populacji?

*https://en.wikipedia.org/wiki/Stochastic_matrix

**<https://www.youtube.com/watch?v=nnssRe5DewE>

***https://en.wikipedia.org/wiki/Absorbing_Markov_chain

Przykład analizy teoretycznej AG

Stany i przejścia

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

Rozważamy populację m osobników o długości k . Ile jest różnych stanów populacji? Ile jest stanów, które są atraktorami?

*https://en.wikipedia.org/wiki/Stochastic_matrix

**<https://www.youtube.com/watch?v=nnssRe5DewE>

***https://en.wikipedia.org/wiki/Absorbing_Markov_chain

Przykład analizy teoretycznej AG

Stany i przejścia

Wstęp

Problemy zawodne

Epistaza

Ocena AE

Neutralność

Dryf genetyczny

Brak mutacji

Zadanie

Bibliografia

Rozważamy populację m osobników o długości k . Ile jest różnych stanów populacji? Ile jest stanów, które są atraktorami?

Użyjemy łańcuchów Markowa; niech π – stan, P – macierz prawdopodobieństw przejścia*. Jest 2^{mk} różnych stanów. Rozkład prawdopodobieństwa stanów, w których możemy wylądować po n krokach** od stanu π to π -ty wiersz macierzy P^n – zapiszmy go jako πP^n . Ponieważ nie ma mutacji, pewne stany są absorbujące (nie ma z nich wyjścia) – wszystkie osobniki są równe; takich stanów jest $a = 2^k$. Zatem macierz P możemy wyrazić jako***

$$P = \begin{bmatrix} I_a & 0 \\ R & Q \end{bmatrix}$$

gdzie I_a to macierz $a \times a$ wypełniona zerami oprócz jedynek po przekątnej, R to podmacierz $t \times a$ opisująca przejścia do stanu absorbującego, Q to podmacierz $t \times t$ opisująca przejścia do stanów nie-absorbujących; $t = 2^{mk} - a$.

*https://en.wikipedia.org/wiki/Stochastic_matrix

**<https://www.youtube.com/watch?v=nnssRe5DewE>

***https://en.wikipedia.org/wiki/Absorbing_Markov_chain

Przykład analizy teoretycznej AG

Agregacja prawdopodobieństw przejścia między stanami

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

Dla n kroków będziemy mieli

$$P^n = \begin{bmatrix} I_a & 0 \\ N_n R & Q^n \end{bmatrix}$$

gdzie $N_n = I_t + Q + Q^2 + Q^3 + \dots + Q^{n-1}$, I_t to macierz $t \times t$ wypełniona zerami oprócz jedynek po przekątnej.

Granicznie

$$\lim_{n \rightarrow \infty} P^n = \begin{bmatrix} I_a & 0 \\ (I_t - Q)^{-1} R & 0 \end{bmatrix}.$$

Przykład analizy teoretycznej AG

Prawdopodobieństwo absorpcji

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

A zatem, co można łatwo zobaczyć **mnożąc przykładowe macierze** w numpy czy nawet w arkuszu kalkulacyjnym, nasz algorytm zaczynając ze stanu nie-absorbującego (prawdopodobieństwa $(I_t - Q)^{-1}R$) wyląduje na pewno w jakimś stanie absorbującym i w nim pozostanie (I_a). Obliczmy prawdopodobieństwo trafienia do stanu absorbującego [Fog00, str. 105]. Niech $\Gamma = \{0, 1\}$. Po n iteracjach, nasz algorytm znajdzie się w stanie γ , $\gamma \in (\Gamma^k)^m$:

$$\Pr(\gamma \in A) = \sum_{i=1}^a (\pi^* P^n)_i = \sum_{i=1}^a \left(\pi^* \begin{bmatrix} I_a \\ N_n R \end{bmatrix} \right)_i$$

gdzie $(\cdot)_i$ oznacza i -ty element poziomego wektora, A to zbiór wszystkich stanów absorbujących, a π^* to wektor poziomy opisujący prawdopodobieństwa rozpoczęcia algorytmu w każdym stanie populacji. Graniczne prawdopodobieństwo absorpcji

$$\lim_{n \rightarrow \infty} \sum_{i=1}^a \left(\pi^* \begin{bmatrix} I_a \\ N_n R \end{bmatrix} \right)_i = \sum_{i=1}^a \left(\pi^* \begin{bmatrix} I_a \\ (I - Q)^{-1} R \end{bmatrix} \right)_i = 1.$$

Zadanie dla chętnych

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

Stosujemy pokoleniowy algorytm ewolucyjny, populacja n osobników, g z n osobników jest dobrych, $n - g$ jest złych. Wykorzystujemy selekcję turniejową (wielkość turnieju k) ze zwracaniem, w której jeśli osobnik dobry spotka się z złym, wygrywa dobry.

Założ, że krzyżowanie i mutacja nie zmieniają jakości osobnika (dobry pozostaje dobrym, zły pozostaje złym).

- 1 Zakładając, że w populacji jest połowa osobników dobrych ($g = \frac{n}{2}$), jak zmieni się ta proporcja w kolejnym pokoleniu (tzn. po jednej selekcji)?
- 2 Zakładając, że wszystkie osobniki podlegają mutacji i mutacja nigdy nie polepsza złego osobnika, ale pogarsza osobnika dobrego z prawdopodobieństwem m , jakie musi być co najmniej g , żeby liczba dobrych w populacji nie spadała?

Bibliografia I

Wstęp

Problemy
zawodne

Epistaza

Ocena AE

Neutralność

Dryf
genetyczny

Brak mutacji

Zadanie

Bibliografia

- [Dav90] Yuval Davidor. "Epistasis variance: Suitability of a representation to genetic algorithms". W: *Complex Systems* 4.4 (1990), s. 369–383. URL: <http://www.complex-systems.com/pdf/04-4-1.pdf>.
- [Fog00] David B. Fogel. *Evolutionary computation: toward a new philosophy of machine intelligence*. 2 wyd. IEEE Press, 2000.
- [Gea+02] Nicholas Geard i in. "A comparison of neutral landscapes – NK, NKp and NKq". W: *Proceedings of the 2002 Congress on Evolutionary Computation, CEC'02*. T. 1. IEEE. 2002, s. 205–210. URL: <https://eprints.soton.ac.uk/264208/1/cec1-comp.pdf>.
- [Gol03] David Edward Goldberg. *Algorytmy genetyczne i ich zastosowania*. Wydawnictwa Naukowo-Techniczne, 2003.
- [NK00] B. Naudts i L. Kallel. "A comparison of predictive measures of problem difficulty in evolutionary algorithms". W: *IEEE Transactions on Evolutionary Computation* 4.1 (2000), s. 1–15. DOI: [10.1109/4235.843491](https://doi.org/10.1109/4235.843491).
- [RW95] Colin R. Reeves i Christine C. Wright. "Epistasis in Genetic Algorithms: An Experimental Design Perspective". W: *ICGA*. 1995, s. 217–224. URL: https://www.researchgate.net/profile/Christine_Wright5/publication/2504851_Epistasis_in_Genetic_Algorithms_An_Experimental_Design_Perspective/links/0912f50bf2d3d75799000000.pdf.



**Fundusze
Europejskie**
Polska Cyfrowa



**Rzeczpospolita
Polska**

Unia Europejska
Europejski Fundusz
Rozwoju Regionalnego

