

Algorytmy genetyczne w interpolacji wielomianowej

(seminarium robocze)

Seminarium Metod
Inteligencji Obliczeniowej
Warszawa 22 II 2006
mgr inż. Marcin Borkowski

Plan:

- Przypomnienie algorytmu niszowego
- Starsze wyniki
- Szereg logistyczny
- Nowe wzorce
- Nowe wyniki

Przypomnienie

- Stała populacja
- Kodowanie binarne
- Nabór:

Osobniki w populacji są sortowane względem ich przystosowania. Osobniki najslabsze są następnie zastępowane przez wyniki operatorów genetycznych.

Pozwala to najlepszym osobnikom unikać “śmierci” tak długo aż zostaną wyparte przez osobniki jeszcze lepsze.

Wyniki operatorów są umieszczane w tymczasowej kopii populacji co daje równe szanse także osobnikom spisanym na straty

Przypomnienie

- Mutacja sterowana dwoma parametrami:
-

Każdy osobnik podlega mutacji z zadany
prawdopodobieństwem mutacji struktury pms .

Każdy bit osobnika wylosowanego do mutacji podlega mutacji
(inwersja) z prawdopodobieństwem pm .

Pm maleje w kolejnych generacjach aż do granicy jednej mutacji
na osobnika

Wynik mutacji wypiera z populacji osobnika najslabszego z
jeszcze niepodmienionych

Przypomnienie

- Krzyżowanie
jednopunktowe:

Wyniki krzyżowania, podobnie jak mutacji, zastępują najsłabsze osobniki w kopii populacji

- Nabór nie namnaża
super osobników,
dłuższe przetrwanie w
populacji gwarantuje
ich pozycja w
rankingu

Przypomnienie

- Nie ma klonowania, unika się wielu osobników tej samej postaci
- Rozwiązań jest kilka, są to najlepsze optima w przeszukiwanej przestrzeni. Ilość tych optimów zależy od rozmiaru populacji

Przypomnienie

- Funkcja oceny:

Ocena rozbita jest na dwa etapy, poza zwyczajową ewaluacją osobnika następuje redukcja oceny na skutek działania czynnika ścisku.

Ścisk wyznaczany jest dopiero gdy znane są wszystkie funkcje oceny

Dla problemów reprezentowanych seriami wylicza się serię ocen cząstkowych. Do oceny końcowej brana jest wartość uśredniona

Przypomnienie

- Ścisk:

W zamyśle ma działać jako narzędzie limitujące ilość podobnych do siebie osobników zamieszkujących tą samą niszę poprzez redukcowanie ich wartości dopasowania.

Nie wszystkie osobniki w niszy będą tak samo zredukowane - redukcja jest proporcjonalna do dotychczasowej wartości dopasowania.

Poza numeryczną wartością dopasowania pod uwagę brane są także inne czynniki np.: skuteczność rozwiązania, długość itp

Osobniki są podobne wtedy, gdy dla tych samych zadanych problemów dają podobne wyniki (podobna fizyczna budowa nie wystarcza)

Przypomnienie

- Ściśk zapewnia rozłączne pokrycie, osobniki należące do więcej niż jednej niszy “cierpią” na większy ścisk.
- Sprzeczności występują tylko w problemach prezentowanych seriami

Przypomnienie

- Działa teza o cegiełkach, wymaga modyfikacji “dowodu”
- Populacja ma nieustannie poszukiwać lepszego zestawu optimów
- Za różnorodność populacji odpowiada mutacja i ścisk

Przypomnienie

- Algorytm wychodzi od losowej populacji
- Po osiągnięciu pierwszych wyników zawsze można oczekiwać lepszych lub zmienić problem i pozwolić populacji przemigrować do nowych nisz

Przypomnienie

- W jednym przebiegu można szukać rozwiązania (rozwiązań) wspólnego dla serii problemów

Starsze wyniki

- Testowano:
 - Interpolacje
 - Predykcje
- Proste szeregi czasowe
 - Schodki
 - Sinus

Starsze wyniki

- Postać rozwiązania:
 - Wzorzec ma postać:
 $(v_1, 0)(v_2, 0_2) \dots (v_n, 0_n)$
 - Odstęp pomiędzy węzłami jest limitowany
 - Dane wejściowe i wartości v_i są skalowane do przedziału $\langle -1, 1 \rangle$
 - Uzupełnianie wzorca nie jest częścią algorytmu, ale oba procesy można przeplatać dość swobodnie

Starsze wyniki

- Każdy wzorzec/osobnik sprawdza się tylko w pewnych określonych warunkach, stopień zgodności jest reprezentowany współczynnikiem zaufania
- Do zadania można włączyć liniowe skalowanie wzorców co rozszerza ich użyteczność
- Do rozwiązania są brane tylko wzorce o dużym współczynniku zaufania i wysokiej wartości dopasowania
- Możliwe jest kilka alternatywnych rozwiązań lub też kompletny brak rozwiązania

Starsze wyniki

- Oceny dopasowania (fitness)
 - Do wyszukania wzorców wykorzystuje się dane uczące (znaną część szeregu).
 - Jedna znana dana jest z szeregu usuwana, a następnie uzupełniana przez wzorce
 - Jeśli predykcja jest zgodna z przewidywaniem - ocena cząstkowa jest wysoka, jeśli nie i jednocześnie współczynnik zaufania jest wysoki – ocena cząstkowa jest niska
 - Proces ten jest powtarzany dla wszystkich znanych danych w szeregu
 - Finalna ocena jest średnią z najlepszych ocen cząstkowych

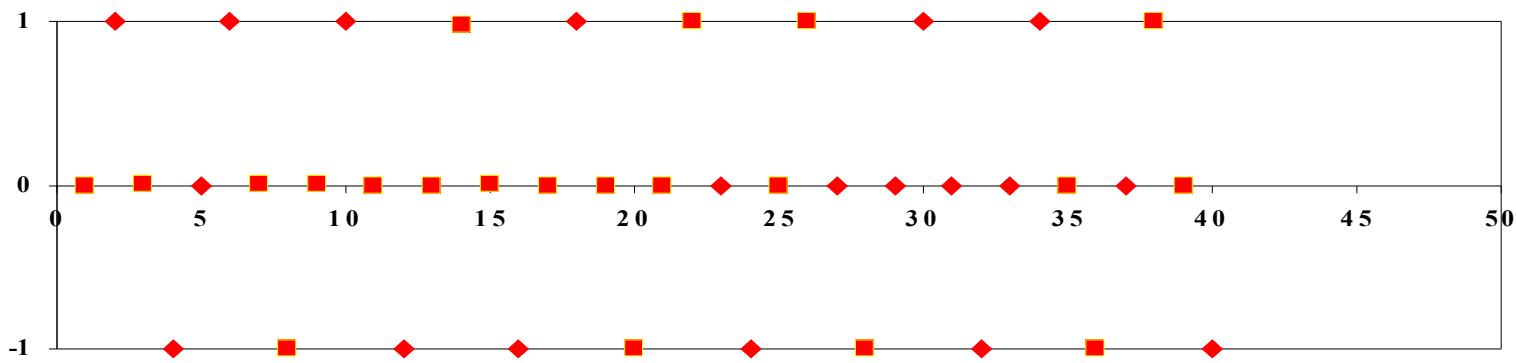
Starsze wyniki

Parametry procesu

- ilość bitów wartości w węźle 12
- maksymalny dystans między węzłami 4
- ilość węzłów we wzorcu <3,4>
- prawdopodobieństwo krzyżowania 45%
- prawdopodobieństwo mutacji struktury 50%
- rozmiar populacji 70
- Początkowe generacje MAG <300,3000>
- Wtórne kroki MAG <100,300>
- Oczekiwany minimalny poziom zaufania <75,95>

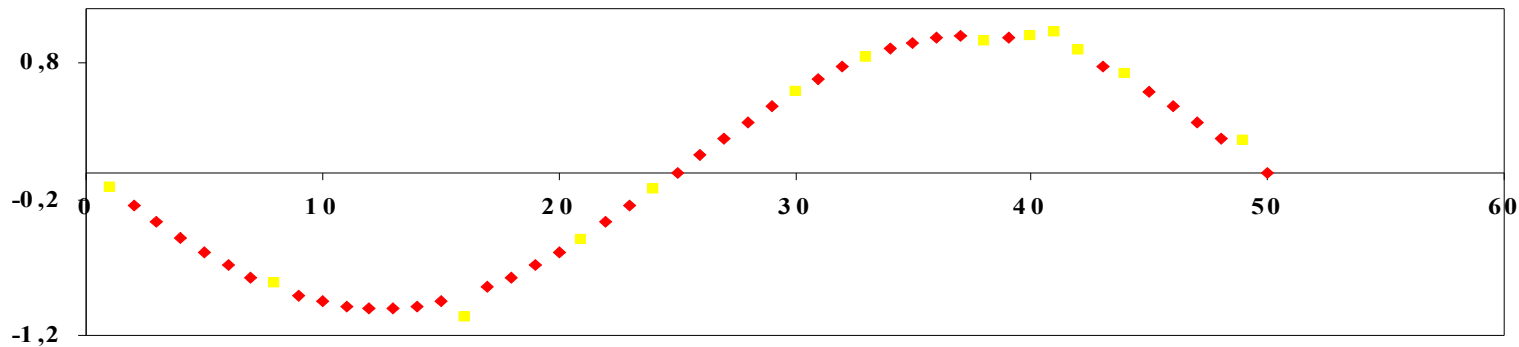
Starsze wyniki

- Odtwarzanie prostego szeregu (52% braków - kwadraty)
- Najlepsze wzorce bez skalowania:
 - $(1.00,0)(0.00,1)(-1.00,2)(1.00,4)$
 - $(-1.00,0)(0.00,1)(1.00,2)(-1.00,4)$
- Najlepsze wzorce ze skalowaniem:
 - $(0.768,0)(0.768,4)(0.768,8)$



Starsze wyniki

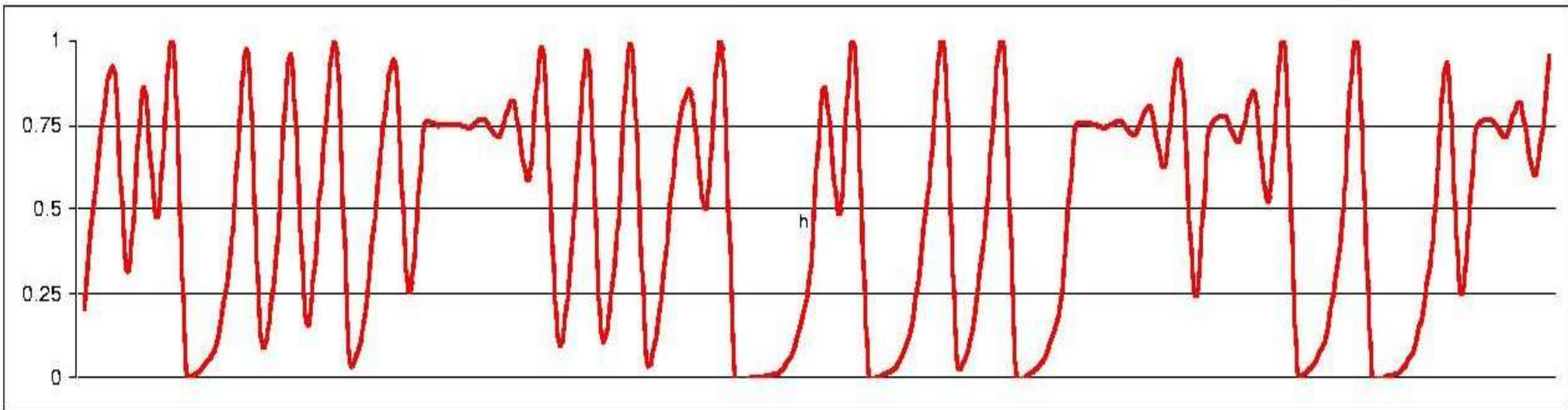
- Sinus – 26 % braków
- Najlepsze wzorce ze skalowaniem:
(-0.75,0)(-0.55,1)(-0.35,2)
(0.86,0)(0.72,1)(0.59,2)
(0.89,0)(-0.07,1)(-0.99,2)
(-1.00,0)(0.03,1)(0.99,2)



Szereg logistyczny

- Równanie logistyczne – system chaotyczny
- $X_{n+1} = R * X_n (1 - X_n)$
- R – stałe
- Wyrazy utrzymują się w przedziale $[0,1]$ dla:
 - X_0 w przedziale $[0,1]$
 - R w przedziale $(0,4]$
- Użyty szereg $x_0 = 0.2027$ $R = 4$

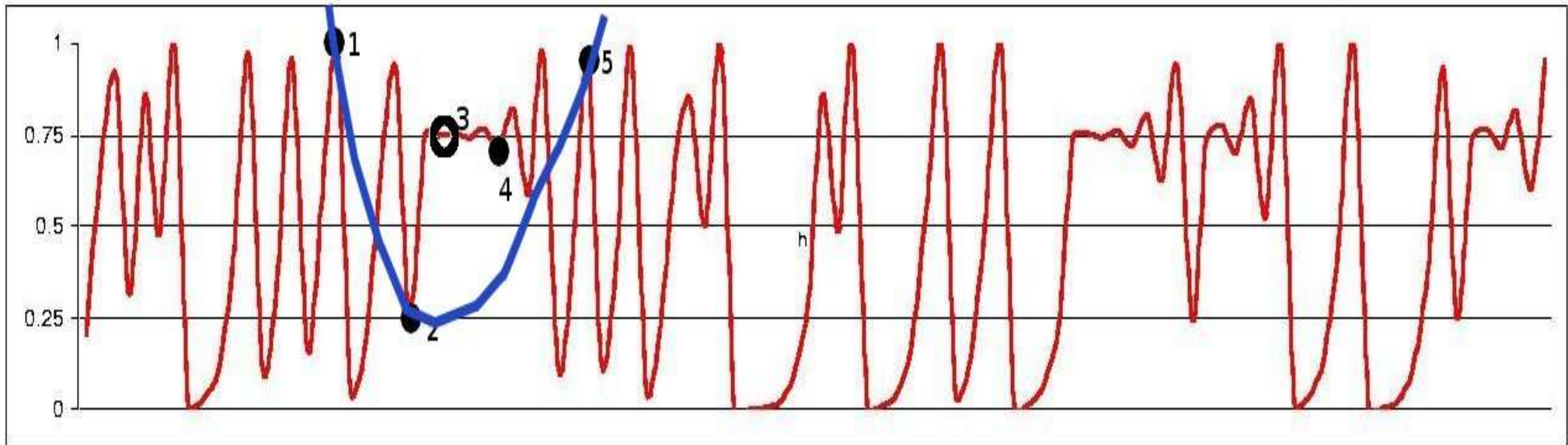
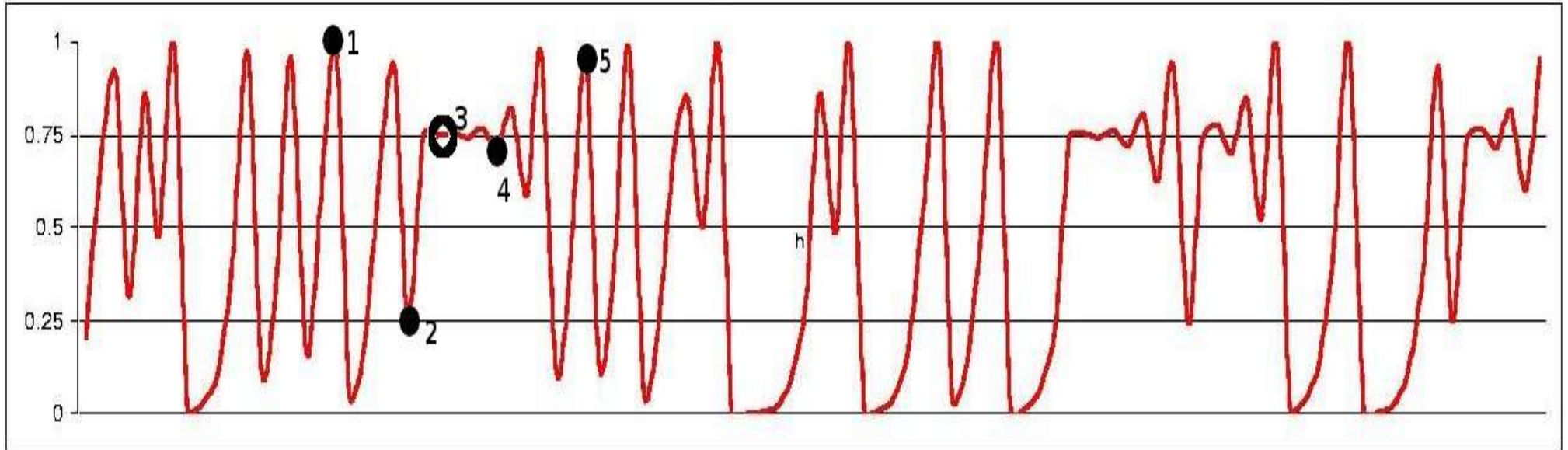
Przykład



Nowe wzorce

- Wielomiany n-tego rzędu:
 - $a_n X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0$
 - Konstrukcja wielomianu na podstawie znanych wartości, konieczne $n+1$ węzłów
 - Weryfikacja wyniku dodatkowy $n+2$ węzeł
 - AG wyszukuje które węzły ze znanych danych użyć do konstrukcji wielomianu
 - Wynik AG to lista przesunięć węzłów względem wartości interpolowanej np:

Wielomian przykład dopasowania



Wielomiany cd.

- Wiarygodność wzorca wyznaczmy sprawdzając czy wybrane węzły należą do jednego wielomianu (w przykładzie 1 2 4 5) poprzez konstrukcję i porównanie wielomianów zbudowanych na $n-1$ węzłach (135 125 245 itd.)
- Przy okazji wielokrotnie wyznaczamy też brakujący węzeł, odchylenie standardowe w tym zbiorze wartości powinno być małe i wpływa na wiarygodność wzorca.

Testy

- Z 40 wyrazowego szeregu logistycznego usuwamy jeden element
- Wykonujemy 10 testów z 1 brakiem
- Powtarzamy test dla wszystkich 40 wartości
- Powtarzamy cały test dla wielomianów od stopnia 2 do 7
- Razem $6 \cdot 40 \cdot 10 = 2400$ testów

Parametry testów

Parametry procesu

- maksymalny dystans między węzłami 4
- ilość węzłów we wzorcu 5-9
- prawdopodobieństwo krzyżowania 35%
- prawdopodobieństwo mutacji struktury 50%
- rozmiar populacji 100
- Początkowe generacje 2000
- Wtórne kroki 0
- Oczekiwany minimalny poziom zaufania 75%

Wyniki testów

- Pojedynczy test
- Procentowo
- Poprawy
- Najlepsze predykcje
- Aktywność populacji

To już wszystko!

Dziękuję za uwagę,
proszę o pytania i komentarze.

Marcin Borkowski
marcinbo@mini.pw.edu.pl