

www.kwmimkm.polisi.pl

OBLICZENIA EWOLUCYJNE

wykład 5

3

www.kwmimkm.polisi.pl

FUNKCJE TESTOWE

F. RASTRIGINA

$-1 \leq x_i \leq 1$

$$f(x) = \sum_{i=1}^n [x_i^2 - \cos(18x_i)]$$

minimum globalne:
 $x = 0, f(x) = -n$
 n - l. zmiennych)

2

www.kwmimkm.polisi.pl

F. ACKLEYA

$$f(x) = -20 \exp\left(-0.2 \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}\right) - \exp\left(\frac{1}{n} \sum_{i=1}^n \cos(2\pi x_i)\right) + 20 + e$$

$-30 \leq x_i \leq 30$

minimum globalne:
 $x = 0, f(x) = 0$

3

www.kwmimkm.polisi.pl

<http://www.ft.utb.cz/people/zelinka/soma/func.html>

Test Functions

Test Functions (some of them are killers for standard methods). If you know some other good function, please let me know. Names behind functions belongs to the researchers who noticed me about them.

- Sphere model, 1st De Jong's function (K. Price, J. Lampinen)
- 2nd De Jong's function (K. Price, J. Lampinen)
- 3rd De Jong's function (K. Price, J. Lampinen)
- Rastrigin's function (K. Price, J. Lampinen)
- Greinwang's function (K. Price, J. Lampinen)
- Stratified V. sine wave function (Ackley)
- Ackley's function (K. Price, J. Lampinen)
- Raman's function (K. Price, J. Lampinen)
- Michalewicz's function (K. Price, J. Lampinen)
- 2nd De Jong's function, Rosenbrock's saddle (K. Price, J. Lampinen)
- 4th De Jong's function (K. Price, J. Lampinen)
- Schwefel's function (K. Price, J. Lampinen)
- Sine envelope sine wave function (K. Price, J. Lampinen)
- Egg holder (K. Price, J. Lampinen)
- Pathological test function (K. Price, J. Lampinen)

4

www.kwmimkm.polisi.pl

TYPOWE OPERATORY KRZYŻOWANIA

5

www.kwmimkm.polisi.pl

Para rodziców – para potomków

Zwykle :
 2 osobniki rodzicielskie - 2 (sprzężone) osobniki potomne.

Pojedynczy osobnik potomny

- wariant dwuosobniczy – para osobników rodzicielskich;
- wariant globalny – jeden wiodący i n pomocniczych osobników rodzicielskich (po jednym dla każdego genu).

Krzyżowanie wieloosobnicze:

- z wieloma osobnikami potomnymi;
- z jednym osobnikiem potomnym.

6

OPERATORY KRZYŻOWANIA WYMIENIAJĄCEGO



7

- Tworzą chromosomy potomne przez składanie ich z wartości genów chromosomów rodzicielskich.
- Mogą być wykorzystywane zarówno przy kodowaniu binarnym, jak i rzeczywistoliczbowym.
- Nie dochodzi do modyfikacji wartości genów zawartych w chromosomach krzyżowanych osobników rodzicielskich
(tylko ich przetasowanie)



8

KRZYŻOWANIE JEDNOPUNKTOWE (proste)

- wybór (z rozkładem jednostajnym) liczby c (punkt rozcięcia) ze zbioru $\{1, 2, \dots, n-1\}$
 n - długość osobnika;

- Podział chromosomów X^1 i X^2 poddawanych krzyżowaniu na dwie części i ich sklejanie:

$$Y = [X^1_1, \dots, X^1_c, X^2_{c+1}, \dots, X^2_n]$$

W wersji z 2 osobnikami potomnymi drugi potomek:

$$Z = [X^2_1, \dots, X^2_c, X^1_{c+1}, \dots, X^1_n]$$



9

X^1	X^2		Y	Z
3.24	2.22		3.24	2.22
-0.22	3.14		-0.22	3.14
1.32	7.72		1.32	7.72
3.22	1.22	c	1.22	3.22
1.20	2.40		2.40	1.20
7.23	4.28		4.28	7.23
-2.21	-2.42		-2.42	-2.21



10

KRZYŻOWANIE DWUPUNKTOWE

- wybór 2 punktów rozcięcia c_1 i c_2 ;
- Podział chromosomów X^1 i X^2 poddawanych krzyżowaniu na 3 części i wymiana środkowej części:

$$Y = [X^1_1, \dots, X^1_{c_1}, X^2_{c_1+1}, \dots, X^2_{c_2}, X^1_{c_2+1}, \dots, X^1_n]$$

W wersji z 2 osobnikami potomnymi drugi potomek:

$$Z = [X^2_1, \dots, X^2_{c_1}, X^1_{c_1+1}, \dots, X^1_{c_2}, X^2_{c_2+1}, \dots, X^2_n]$$

$c_1 = c_2$ – krzyżowanie jednopunktowe



11

X^1	X^2		Y	Z
3.24	2.22		3.24	2.22
-0.22	3.14		-0.22	3.14
1.32	7.72		1.32	7.72
3.22	1.22	c_1	1.22	3.22
1.20	2.40		2.40	1.20
7.23	4.28	c_2	4.28	7.23
-2.21	-2.42		-2.21	-2.42



12

KRZYŻOWANIE WIELOPUNKTOWE...

KRZYŻOWANIE RÓWNOMIERNE

www.kwmimkn.polsl.pl

Chromosom potomny:

jeśli wylosowano liczbę $< p_e$;
w przeciwnym razie.

p_e – parametr krzyżowania (typowo $p_e=0.5$)

W wersji z 2 osobnikami potomnymi drugi potomek:

jeśli $Y_i = X^1_i$;
w przeciwnym razie.



13

$p_e=0.5$

www.kwmimkn.polsl.pl

X^1	X^2	wylosowano	Y	Z
3.24	2.22	0.092699	3.24	2.22
-0.22	3.14	0.158384	-0.22	3.14
1.32	7.72	0.697190	7.72	1.32
3.22	1.22	0.315814	3.22	1.22
1.20	2.40	0.821422	2.40	1.20
7.23	4.28	0.399981	7.23	4.28
-2.21	-2.42	0.428556	-2.21	-2.42



14

KRZYŻOWANIE DIAGONALNE

www.kwmimkn.polsl.pl

- Jest krzyżowaniem wieloosobniczym.
- Tworzy r potomków z r rodziców przy $c = r - 1$ punktach krzyżowania.
- Osobniki potomne powstają w wyniku składania fragmentów kodu po przekątnej.

Dla 3 osobników:

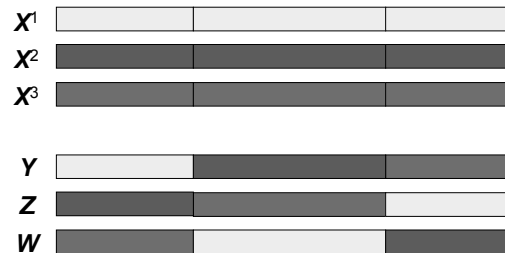
$$Y = [X^1_1, \dots, X^1_{c1}, X^2_{c1+1}, \dots, X^2_{c2}, X^3_{c2+1}, \dots, X^3_n]$$

$$Z = [X^2_1, \dots, X^2_{c1}, X^3_{c1+1}, \dots, X^3_{c2}, X^1_{c2+1}, \dots, X^1_n]$$

$$W = [X^3_1, \dots, X^3_{c1}, X^1_{c1+1}, \dots, X^1_{c2}, X^2_{c2+1}, \dots, X^2_n]$$



15



W wersji 1 potomkiem – tylko potomek Y



16

OPERATORY KRZYŻOWANIA UŚREDNIAJĄCEGO

www.kwmimkn.polsl.pl



17

- Są specyficzne dla kodowania rzeczywistości liczbowego;
- Oddziałują na wartości genów chromosomów poddawanych krzyżowaniu;
- Wartości każdego genu chromosomów potomnych są liczbami zawierającymi się między największą i najmniejszą wartością genu chromosomów rodzicielskich.

www.kwmimkn.polsl.pl



18

KRZYŻOWANIE ARYTMETYCZNE

www.kwmimkm.polssi.pl

- generowanie liczby losowej k z zakresu $(0,1)$ lub jej arbitralny wybór;
- uśrednianie arytmetyczne wartości genów chromosomów rodzicielskich:

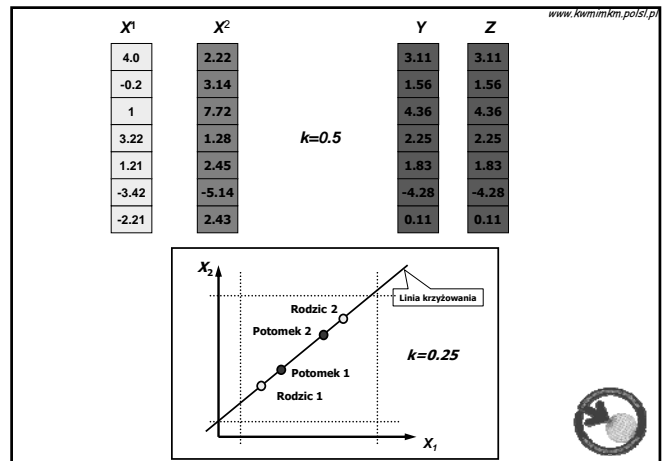
$$Y = X^1 + k(X^2 - X^1)$$

W wersji z 2 osobnikami potomnymi drugi potomek:

$$Z = X^2 + X^1 - Y^1$$



19



KRZYŻOWANIE HEURYSTYCZNE

www.kwmimkm.polssi.pl

- Nie jest krzyżowaniem uśredniającym!
- Generowanie liczby losowej k z zakresu $(0,1)$;
- Tworzy się (maksymalnie) jednego potomka:

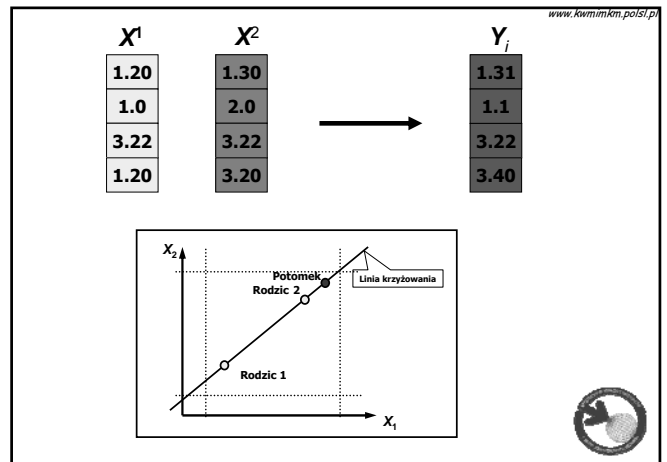
$$Y = k(X^2 - X^1) + X^2$$

przy założeniu, że $f(X^2) \geq f(X^1)$

- Może utworzyć potomka, który nie jest dopuszczalny, wówczas:
 - » generuje się nową liczbę losową i tworzy nowego potomka;
 - » Jeśli po założonej liczbie prób nie utworzono osobnika dopuszczalnego, to nie tworzy się potomka.



21



TYPOWE OPERATORY MUTACJI

www.kwmimkm.polssi.pl



23

MUTACJA RÓWNOMIERNA

www.kwmimkm.polssi.pl

- Losowy wybór genu w chromosomie.
- Przyjęcie przez gen wartości losowej (z rozkładem równomiernym) z zakresu dopuszczalnego dla danej zmiennej:

$$Y = [X_1, \dots, X'_k, \dots, X_n],$$

$$X'_k = \langle \text{left}(k), \text{right}(k) \rangle$$

Szczególnie użyteczna we wczesnej fazie działania AF (gdy pożądane jest szerokie przeszukiwanie obszaru poszukiwań optimum).



24

MUTACJA NIERÓWNOMIERNA

www.kwmimkm.polisi.pl

- Należy do grupy tzw. mutacji ze strojeniem.
- Modyfikacja wartości wybranego genu o wartość pewnej funkcji $\Delta(t,y)$:

$$Y = [X_1, \dots, X'_k, \dots, X_n],$$

gdzie:

$$X'_k = X_k + \Delta(t, \text{right}(k) - X_k) \text{ gdy wylosowano } 0$$

$$X'_k = X_k - \Delta(t, X_k - \text{left}(k)) \text{ gdy wylosowano } 1$$

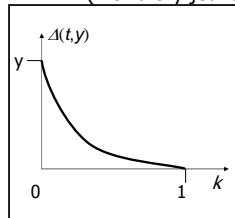


25

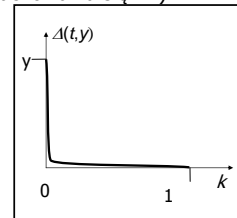
- Funkcja $\Delta(t,y)$ przyjmuje wartości z zakresu $[0,y]$;

- Prawdopodobieństwo, że $\Delta(t,y)$ jest bliskie zero wzrasta ze wzrostem czasu obliczeń

(nie zależy jednak od zachowania się AE).



Początkowa faza obliczeń



Pod koniec działania AE



26

MUTACJA BRZEGOWA

www.kwmimkm.polisi.pl

- Jest odmianą mutacji równomiernej, w której:

$$X'_k = \text{left}(k) \text{ gdy wylosowano } 0$$

$$X'_k = \text{right}(k) \text{ gdy wylosowano } 1$$

- Szczególnie użyteczna, gdy rozwiązanie optymalne leży na brzegu obszaru dopuszczalnego lub bardzo blisko tego brzegu).



27

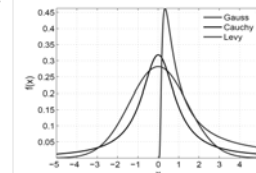
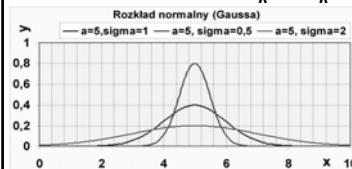
MUTACJA GAUSSOWSKA

www.kwmimkm.polisi.pl

- Przyjęcie przez wylosowany gen wartości losowej (z rozkładem Gaussa) o wartości oczekiwanej równej wartości przed zmianą:

$$Y = [X_1, \dots, X'_k, \dots, X_n],$$

$$X'_k = X_k + N(0, \sigma)$$



28

ISTOTNOŚĆ OPERATORÓW

www.kwmimkm.polisi.pl

- W AG przyjmuje się często, że krzyżowanie jest operatorem pierwszoplanowym, podczas gdy mutacja ma za zadanie zapewnić „dopływ świeżej krwi” do populacji i jest operatorem o mniejszym znaczeniu.
- Wśród badaczy uprawiających programowanie ewolucyjne twierdzi się, że krzyżowanie jest operatorem zbędnym, zaś mutacja jest jedynym mechanizmem przeszukiwania.
- Potrzebny jest co najmniej jeden taki operator genetyczny, który gwarantuje spójność przestrzeni genotypów - operatorem tym jest najczęściej mutacja



29

OPERATORY GENETYCZNE Z ADAPTUJĄCYMI SIĘ PRAWDOPODOBIENSTWAMI

www.kwmimkm.polisi.pl

Użyteczność większości operatorów genetycznych nie jest jednakowa w każdej fazie działania algorytmu.

Mechanizm adaptacji:

- Prawdop. operatorów genetycznych, które częściej prowadzą do lepszych osobników potomnych wzrasta kosztem operatorów mających gorsze wyniki.
- Należy zadbać, aby nie zablokować żadnego z operatorów (prawdopodobieństwa nie powinny osiągać zera)
- Skuteczność działania operatora - oceniana przez zewnętrzny proces monitorujący który steruje prawdop. operatorów.
- Należy uwzględnić w algorytmie korzyści wynikające z działania pewnych operatorów widoczne dopiero po kilku pokoleniach.



30

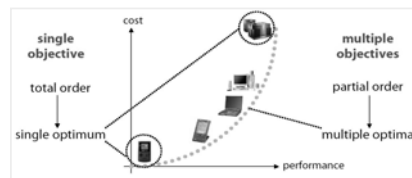
OPTYMALIZACJA WIELO- KRYTERIALNA



31

SFORMUŁOWANIE ZADANIA

- Wiele praktycznych problemów podejmowania decyzji, projektowania itp. trudno jest sformułować jako zadanie optymalizacji funkcji celu zwracającej jedną wartość.
- Zamiast jednego liczbowego kryterium oceny musi się uwzględnić cały ich zbiór.
- Nierzadko jednoczesna minimalizacja tych kryteriów jest niemożliwa z powodu ich wzajemnej sprzeczności.



32

Formalnie:

- Dany jest zbiór (wektor) m funkcji:

$$f(x) = [f_1(x) \dots f_m(x)].$$

- Celem jest jednoczesna minimalizacja wszystkich kryteriów $f_j(x)$.
- Z reguły kryteria nie są ze sobą zgodne (minimalizacja względem jednego z nich może powodować wzrost innych).
- Rozwiązanie x jest zdominowane wtedy i tylko wtedy gdy istnieje dopuszczalne rozwiązanie y nie gorsze od x (dla wszystkich kryteriów).
- Jeśli rozwiązanie nie jest zdominowane przez żadne inne rozwiązanie dopuszczalne to nazywamy je rozwiązaniem niezdominowanym lub rozwiązaniem paretooptymalnym.



33

Przykład:

Gra na giełdzie papierów wartościowych

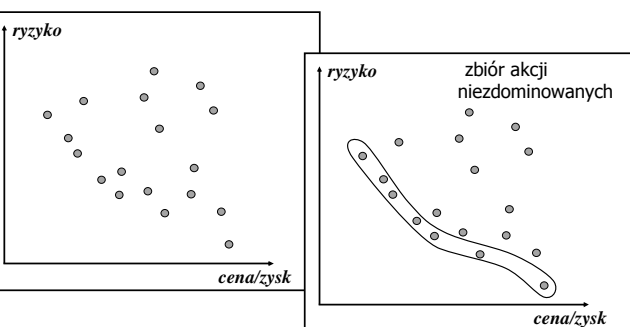
- Cel: kupowanie akcji charakteryzujących się minimalnym ryzykiem i jak maksymalnym zyskiem.
- Z reguły nie ma takich akcji, które spełniałyby oba kryteria jednocześnie.
- Należy odrzucić akcje, dla których istnieje co najmniej jedna inna charakteryzująca się jednocześnie mniejszym ryzykiem i wyższą rentownością.
- Akcje, które pozostały po takiej selekcji, są równoprawnymi rozwiązaniami zadania jednoczesnej minimalizacji ryzyka i współczynnika cena/zysk.



34

Cel:

Investowanie w akcje o jak najmniejszym ryzyku i jak najmniejszej wartości współczynnika cena/zysk.



35

Zadanie optymalizacji wielokryterialnej polega na poszukiwaniu zbioru P punktów niezdominowanych.

Szczególne przypadki:

- Poszukiwanie niezdominowanego punktu;
- Poszukiwanie zbioru punktów niezdominowanych:
 - dużo trudniejsze i w ogólnym przypadku niemożliwe do rozwiązania;
 - często spotyka się uproszczenie, polegające na znalezieniu możliwie największej ich liczby).



36

www.kwmimkm.polisi.pl

Metody optymalizacji wielokryterialnej

A-priori:


- decyzja podejmowana przed rozpoczęciem optymalizacji (pojedynczy cel uzyskany a priori).
„Zwykłe” metody optymalizacji mogą być użyte....

Interaktywne:

- decyzja podejmowana podczas przeszukiwania.
Wymagana interakcja z użytkownikiem...

A-posteriori:

- znalezienie zbioru rozwiązań niezdominowanych, następnie podjęcie decyzji co do wyboru rozwiązania.
Konieczne metody optymalizacji wielokryterialnej...



37

www.kwmimkm.polisi.pl

Metody optymalizacji wielokryterialnej

	A-PRORI	INTERAKTYWNE	A-POSTERIORI
Zalety	<ul style="list-style-type: none"> • niski koszt obliczeniowy • proste w implementacji • łatwo dostępne oprogramowanie 	<ul style="list-style-type: none"> • decydent nadzoruje proces optymalizacji • decydent „uczy się” problemu 	<ul style="list-style-type: none"> • preferencje są określone po fazie optymalizacji • uzyskujemy więcej, niż 1 rozwiązanie
Wady	<ul style="list-style-type: none"> • wymagana (nierealistycznie) duża wiedza o problemie • wymagana analiza wrażliwości 	<ul style="list-style-type: none"> • niezbędna intensywna interakcja z decydem • wynik zależny od wiedzy decydem 	<ul style="list-style-type: none"> • kosztowne obliczeniowo • potrzebna druga faza dla dokonania wyboru; • oprogramowanie nie jest tak powszechne.



38


www.kwmimkm.polisi.pl

SKALARYZACJA ZAD. WIELOKRYTERIALNEGO

Sprowadzenie do problemu jednokryterialnego poprzez wprowadzenie dodatkowego kryterium, porządkującego punkty niezdominowane.

1. Określenie maksymalnych dopuszczalnych wartości poszczególnych składników

- Problem wielokryterialny sprowadza się do zagadnienia znalezienia dowolnego punktu dopuszczalnego.
- Jedno z kryteriów może być uznane za wiodące.
- Rozwiązania poszukuje się stosując metody właściwe dla problemów z ograniczeniami.



39

www.kwmimkm.polisi.pl

2. Wprowadzenie funkcji agregującej.


- Argumenty – wartości poszczególnych składników wektorowego wskaźnika jakości.
- W wyniku - zagregowany, skalarny wskaźnik jakości $f^s(x)$, będący następnie przedmiotem optymalizacji:

$$f^s(x) = \varphi_s [f_1(x), \dots, f_m(x)]$$

φ_s - funkcja agregująca

Zależnie od funkcji φ_s – różne metody skalaryzacji, np.:

1. ważone sumowanie kryteriów,
2. metoda punktu idealnego,
3. metoda punktu najgorszych oczekiwań.



40


www.kwmimkm.polisi.pl

WAŻONE SUMOWANIE SKŁADNIKÓW WEKTOROWEGO WSKAŹNIKA JAKOŚCI

- Najczęściej stosowana technika skalaryzacji.
- Funkcja agregująca φ_s jest liniowa – w postaci sumy ważonej składników wektorowego wskaźnika jakości.

$$\varphi_s [f_1(x), \dots, f_m(x)] = \sum_{k=1}^m w_k f_k(x)$$

- „Wycena” poszczególnych elementów wektorowego wskaźnika jakości.
- Można zaakceptować pogorszenie jednego z kryteriów o pewną wielkość, jeśli skompensują to korzyści wynikające z poprawy wartości innych kryteriów;
- Wzajemne proporcje korzyści i strat wynikają z przyjętych współczynników wagowych w_k .



41

www.kwmimkm.polisi.pl


METODA PUNKTU IDEALNEGO

- Decydem podaje punkt, zwany idealnym, znajdujący się poza obszarem dopuszczalnym. Punkt ten to idealne wartości wektorowego wskaźnika jakości.
- Zadanie optymalizacji wielokryterialnej sprowadza się do znalezienia rozwiązań niezdominowanych znajdujących się najbliżej punktu idealnego.
- Funkcja agregująca ma postać:

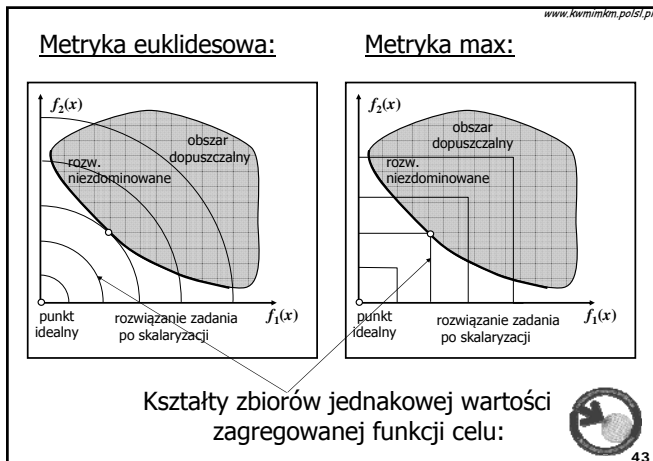
$$\varphi_s [f_1(x), \dots, f_m(x)] = |f(x) - \mathbf{f}^i|$$

$$\min(\varphi_s)$$

- \mathbf{f}^i – punkt idealny
- $|\cdot|$ – norma wektora.



42



www.kwmimkm.polssi.pl

METODA PUNKTU NAJGORSZYCH OCZEKIWAŃ

- Metoda dualna do metody punktu idealnego.
- W metodzie tej maksymalizuje się odległość od tzw. punktu najgorszych oczekiwań.

$$\max(\varphi_s)$$

MINIMA LOKALNE SKALARNEJ FUNKCJI CELU

- Zastosowanie metody skalaryzacji może prowadzić do funkcji celu posiadającej więcej niż jedno minimum.
- Wskazane jest wówczas zastosowanie jednej z metod optymalizacji wielomodalnej.

44

www.kwmimkm.polssi.pl

Ewolucyjne metody optymalizacji wielokryterialnej

- Zwykle bazują na podejściu Pareto;
- Rozwiązania powinny być równomiernie rozmieszczone na froncie Pareto;
- AE przetwarzają jednocześnie wiele rozwiązań, lecz problemem jest globalność selekcji;
- Należy wprowadzić odpowiednie mechanizmy (współczynnik zatłoczenia, niszczenie itp.)
- AE: osobnik A jest lepszy od osobnika B jeśli A ma wyższe przystosowanie.
- tu: osobnik A jest lepszy od osobnika B jeśli go dominuje.

45

www.kwmimkm.polssi.pl

Popularne AE optymalizacji wielokryterialnej:

- VEGA: Vector Evaluated Genetic Algorithm (Schaffer 1985),
- HLGA: Hajela's and Lin's Weighting-based GA (1992),
- FFGA: Fonseca's and Fleming's Multiobjective GA (1993),
- NPGA: The Niche Pareto GA (Horn, Nafpliotis, Goldberg 1994),
- NSGA: The Nondominated Sorting GA (Srinivas, Deb 1994),
- SPEA: The Strength Pareto EA (Zitzler, Thiele 1999).
- SPEA2 (Zitzler, 2001)
- NSGA-II (Deb i inni, 2000)

46

www.kwmimkm.polssi.pl

np. NSGA-II:

- Sortuje osobniki w fronty.
- Stosuje się „crowding distance” jako kryterium dodatkowe.
- Operatory:
 - mutacja równomierna
 - krzyżowanie proste
 - selekcja bazująca na turniejowej

47

www.kwmimkm.polssi.pl

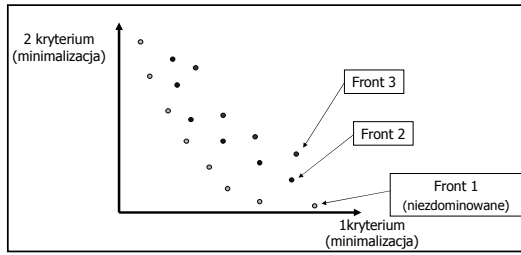
Sortowanie:

Cel: wybrać N „najlepszych” spośród $2 \cdot N$ osobników

48

Podział na fronty:

www.kwmimkn.polsl.pl



„Crowding distance”:

www.kwmimkn.polsl.pl

- Dotyczy osobników należących do tego samego frontu.
- Jest to miara wolnej przestrzeni wokół danego rozwiązania (więcej=lepiej).
- Skutkuje wyborem osobników znajdujących się w mniej zatłoczonych rejonach.

