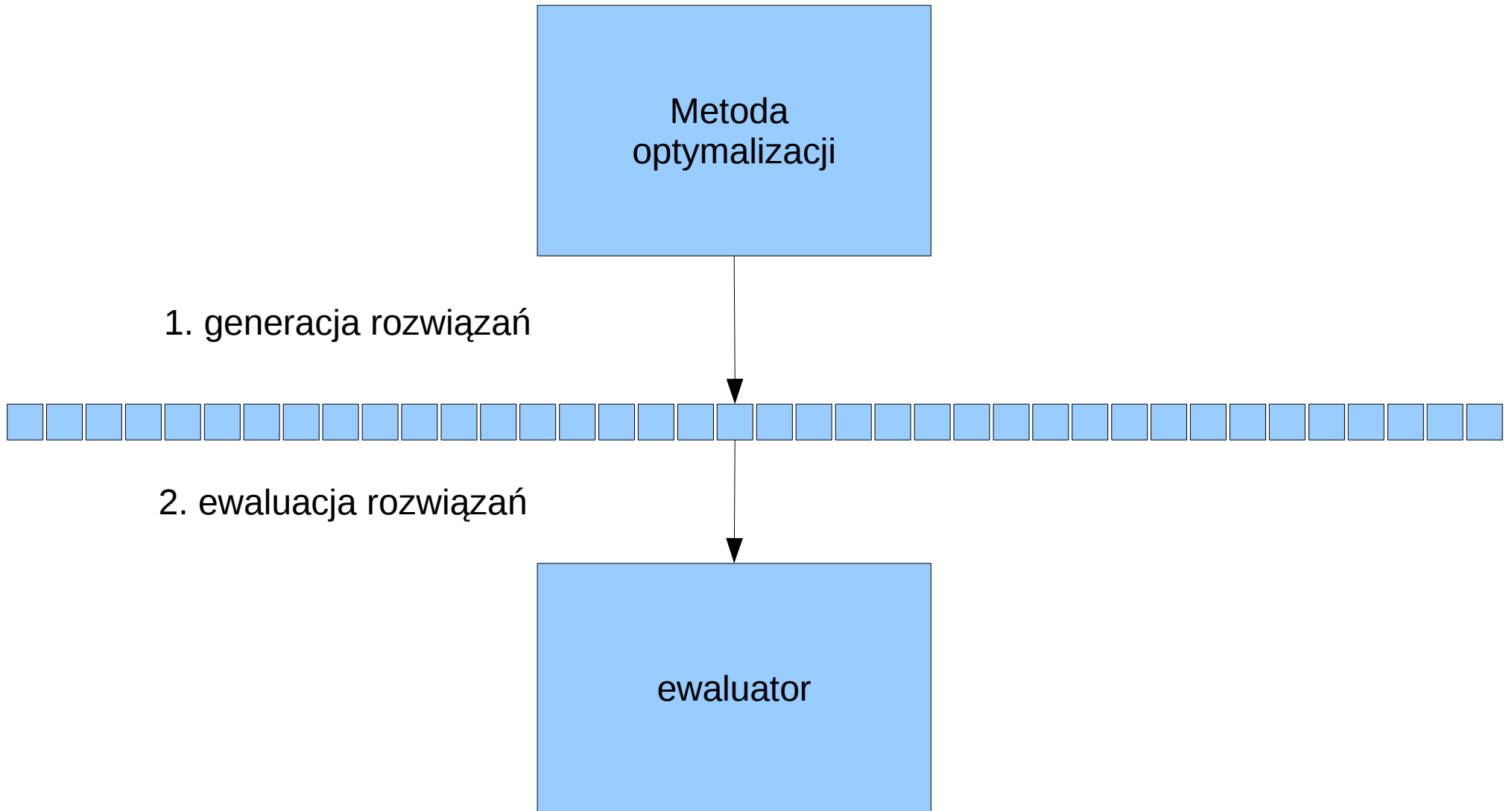


WAE

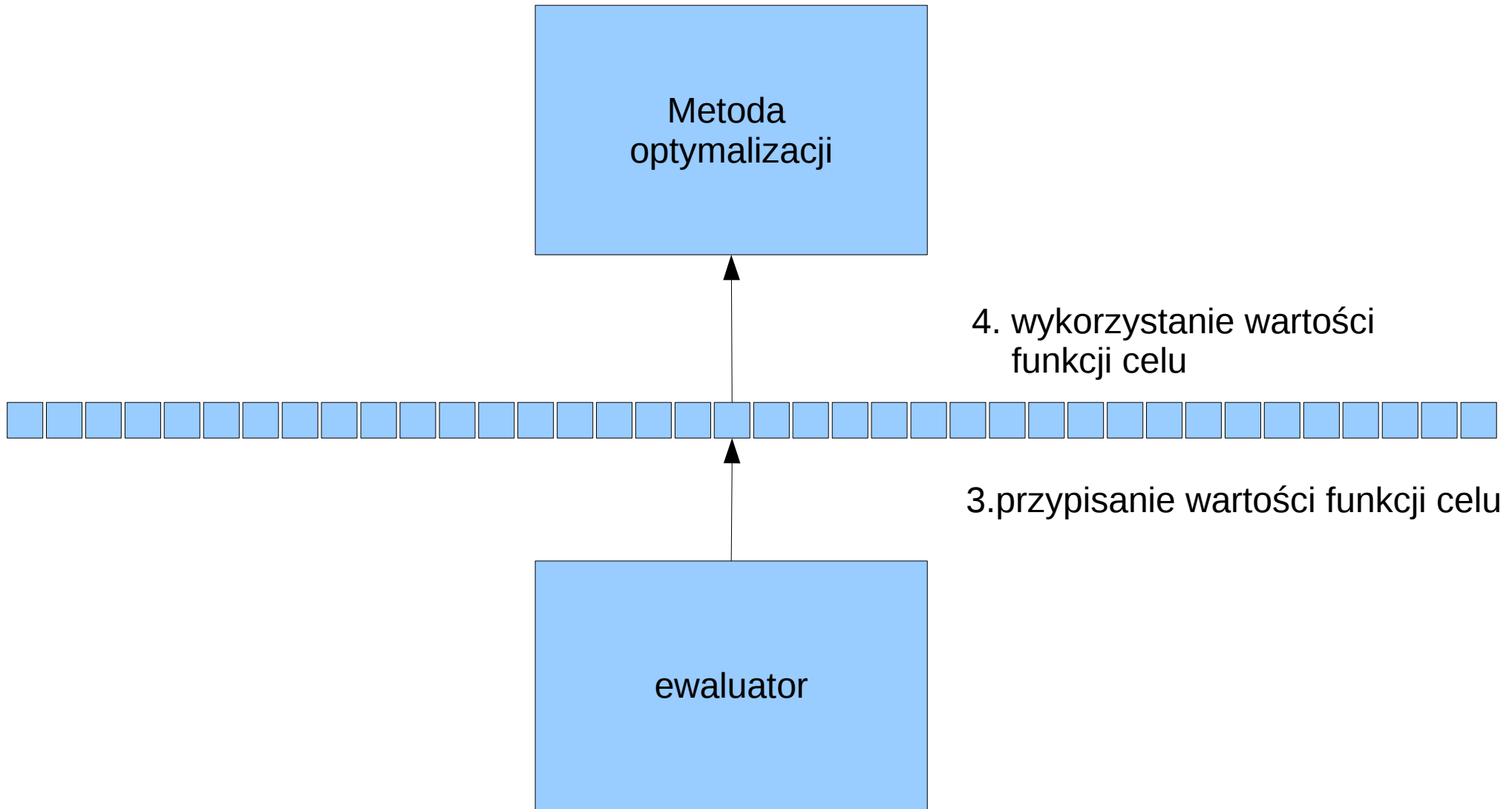
Jarosław Arabas

Testowanie metod optymalizacji

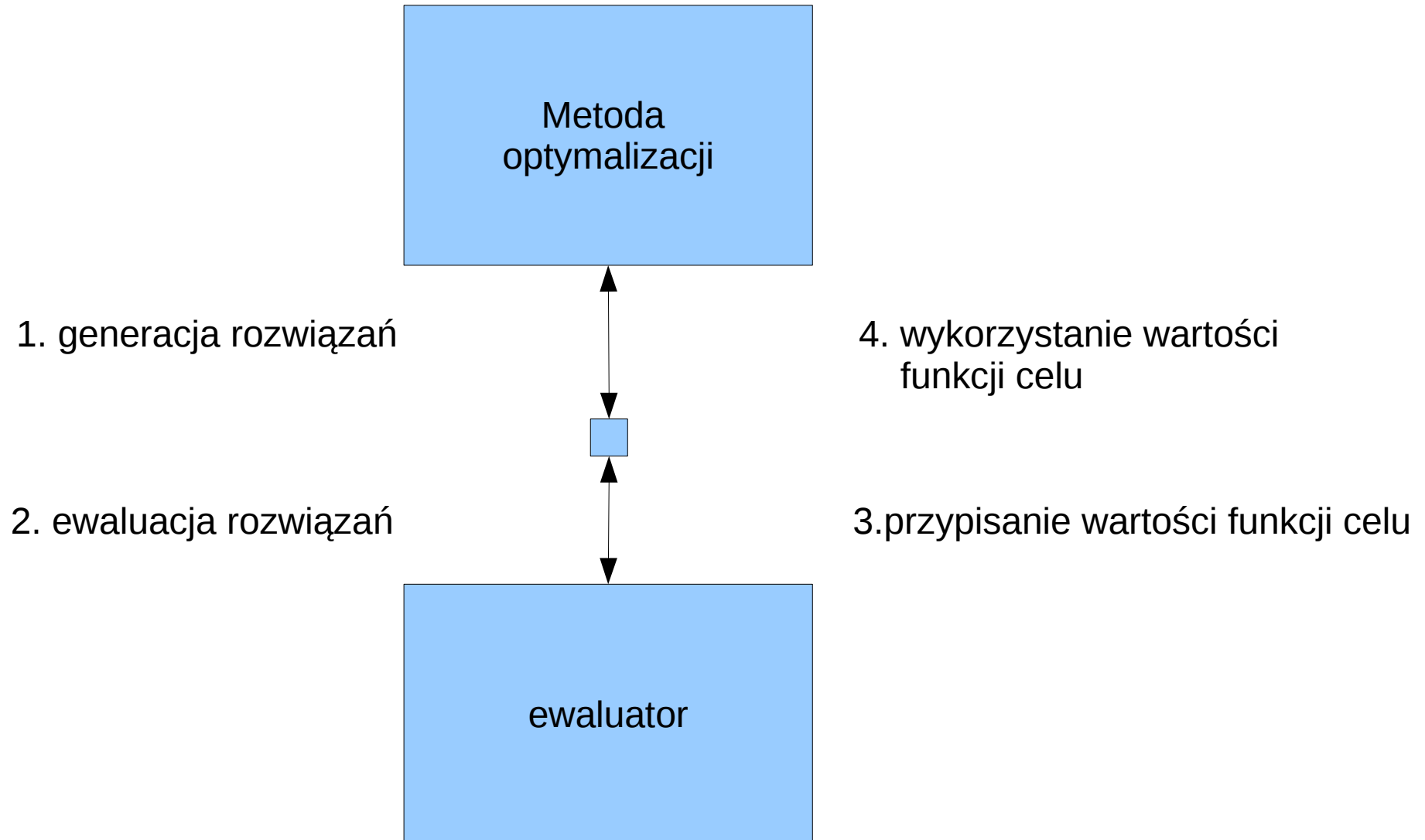
# Metoda optymalizacji jako generator propozycji rozwiązań



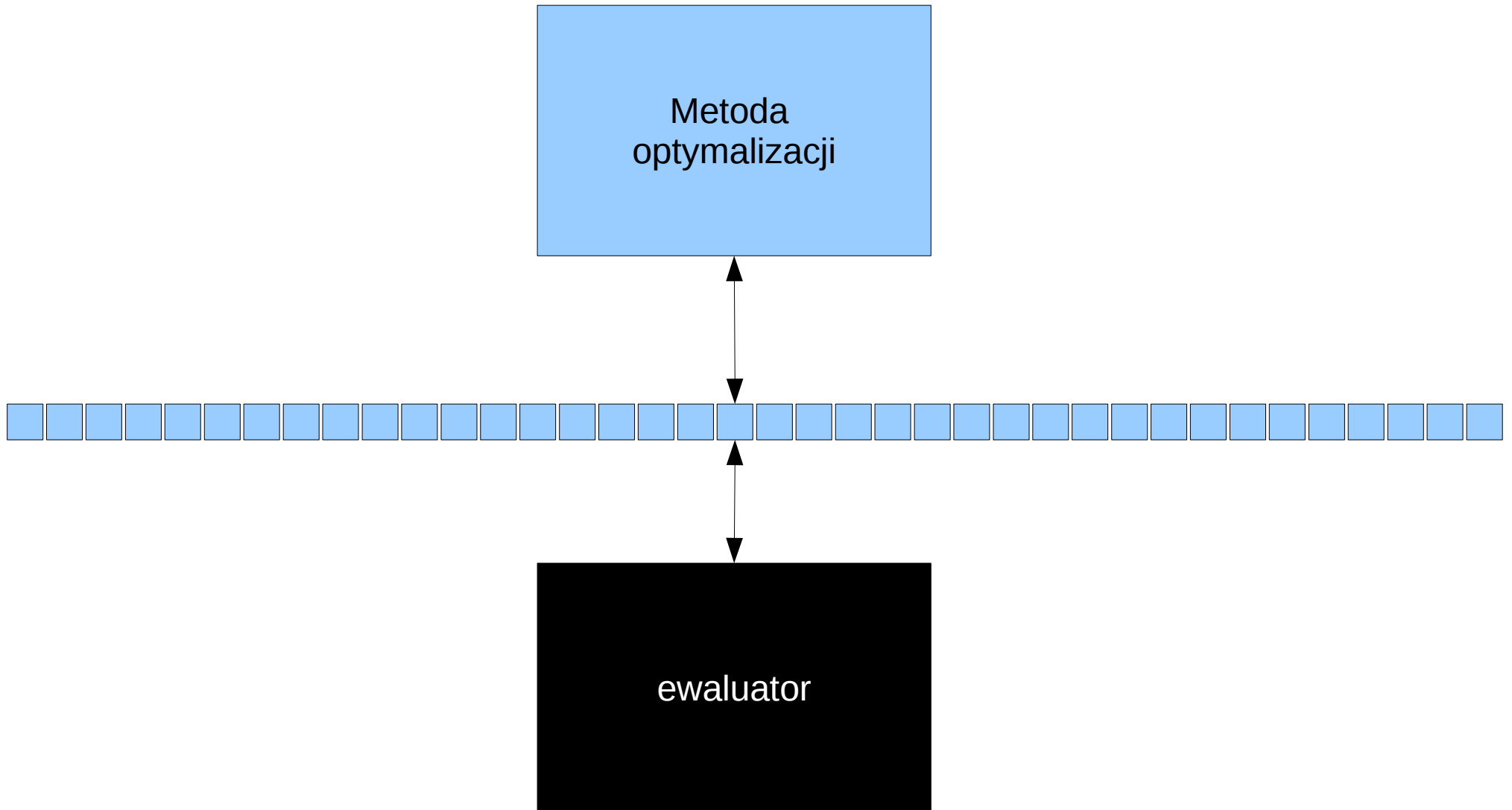
# Metoda optymalizacji jako generator propozycji rozwiązań



# Informacyjne sprzężenie zwrotne optymalizator – ewaluator



# Optymalizacja typu “czarna skrzynka” (*black box*)



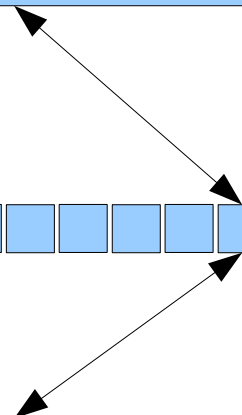
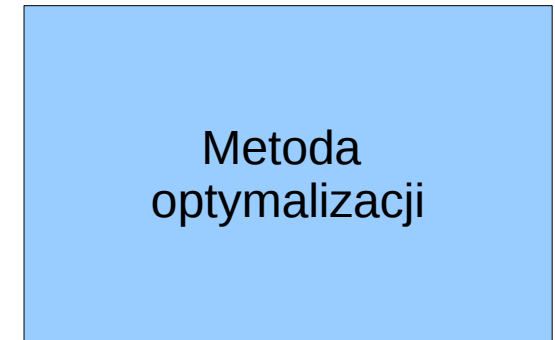
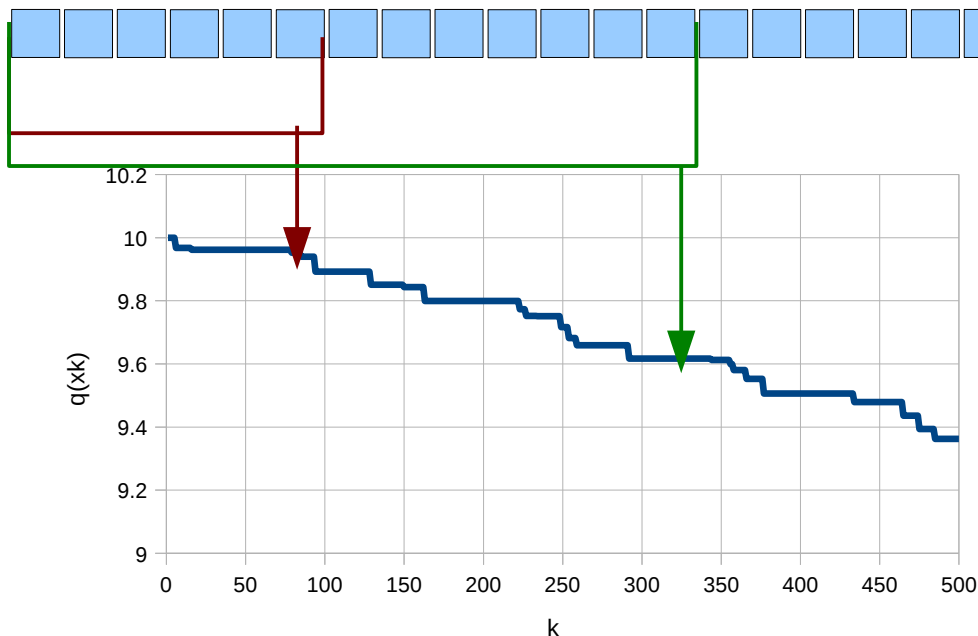
# Testowanie metod optymalizacji z założeniem “czarnej skrzynki”

- Zadania benchmarkowe

CEC 2005 ... 2017

BBOB 2009

- Krzywe “zbieżności” (minimalizacja)



# No Free Lunch Theorem (NFL)



- Nie wiedząc z góry, jakie zadanie optymalizacji przyjdzie nam rozwiązać, nie możemy dobrać efektywnej metody dla niego



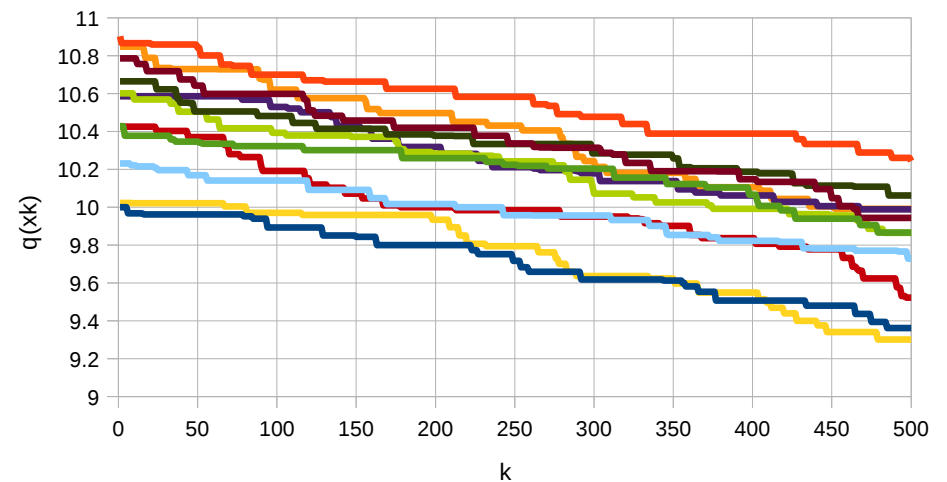
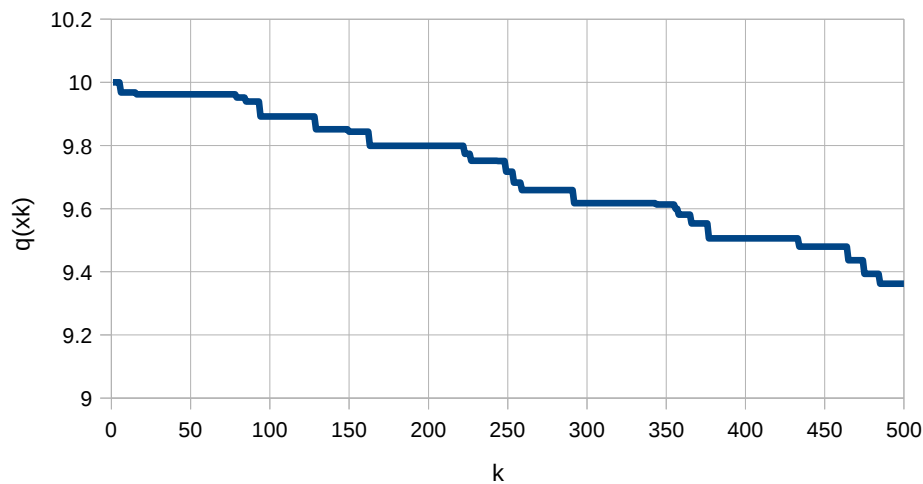
# Testowanie metod optymalizacji z założeniem “czarnej skrzynki”

- Zadania benchmarkowe

CEC 2005 i następne ... CEC2013

BBOB 2009 i 2010

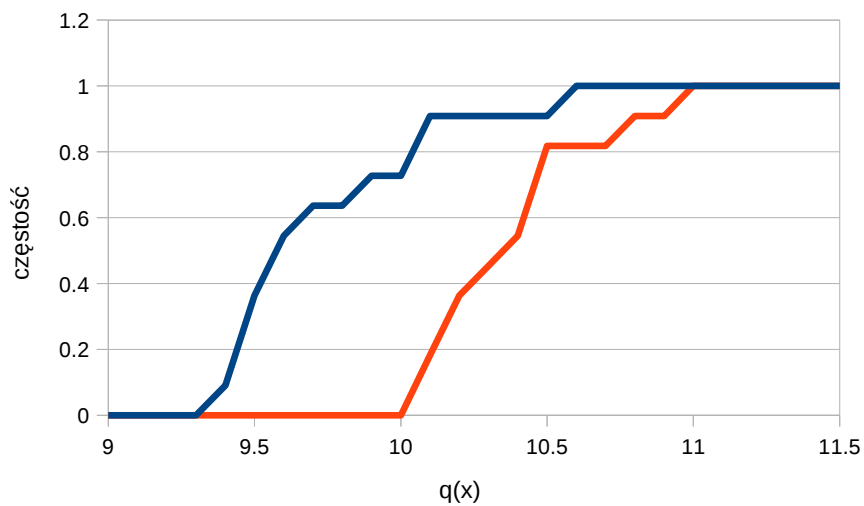
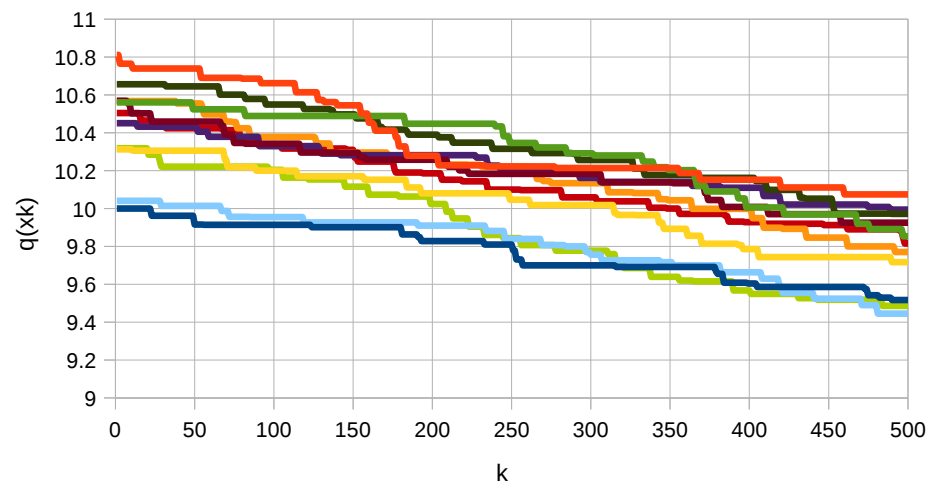
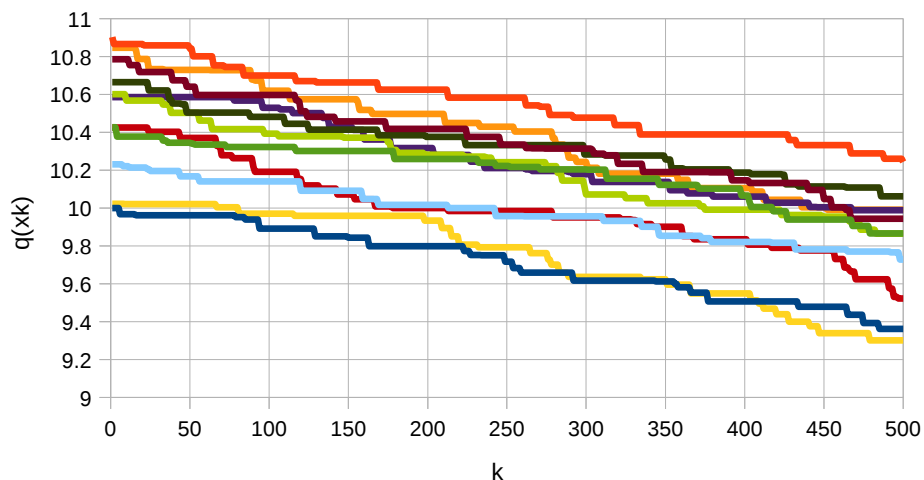
- Krzywe “zbieżności”





# Testowanie metod optymalizacji

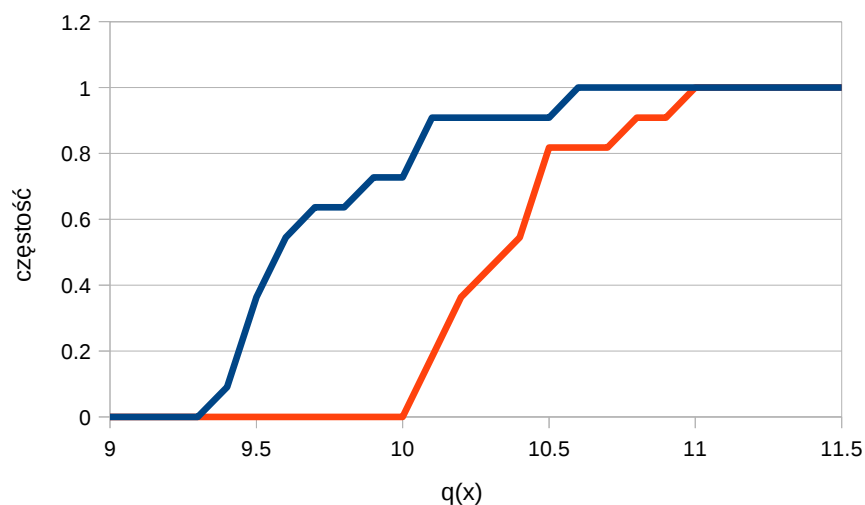
- Porównanie metod losowych



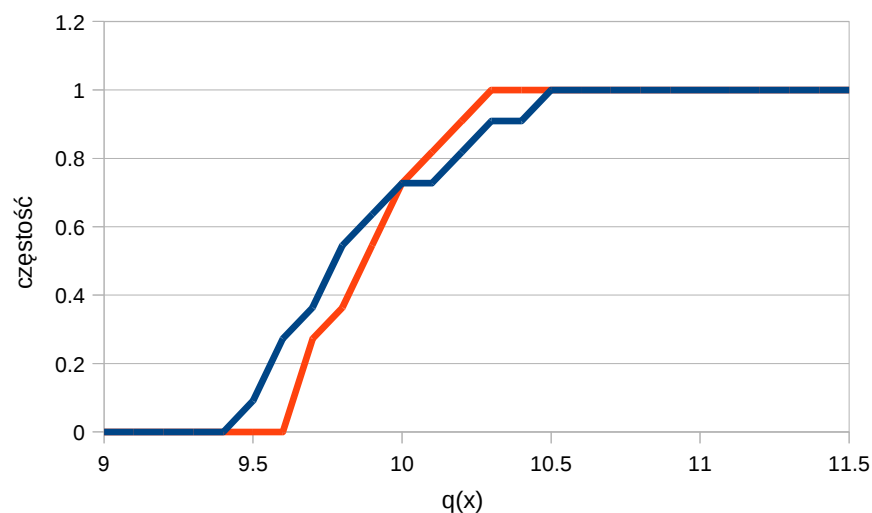
Dystrybuanta empiryczna rozwiązań uzyskanych po założonym czasie dla dwóch algorytmów

# Testowanie metod optymalizacji

- Porównanie algorytmów na podst. dystrybuanty rozwiązań (testy nieparametryczne)



**Algorytm B**  
jest gorszy niż  
**algorytm A**



Nie można stwierdzić,  
że jeden jest lepszy  
od drugiego

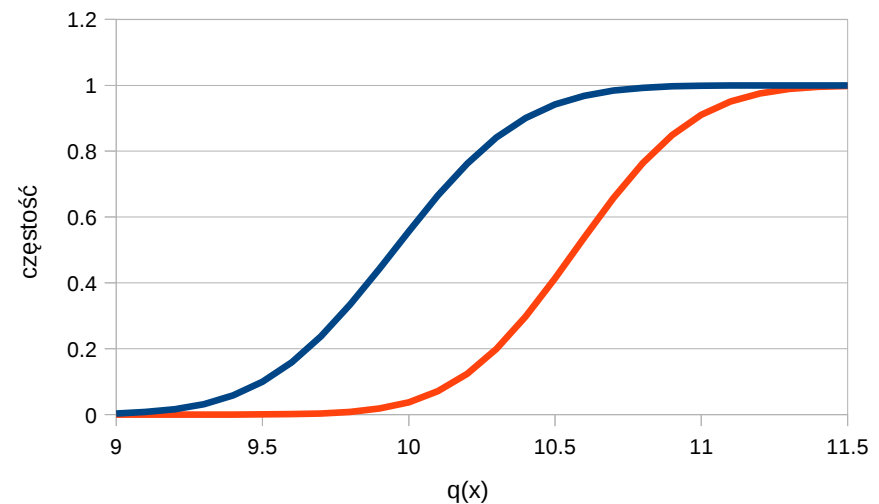
# Testowanie metod optymalizacji

- Porównanie algorytmów na podst. porównania statystyk (testy parametryczne)

10.57 0.32

9.95 0.35

Algorytm B  
jest gorszy niż  
algorytm A



Dystrybuanty rozkładu  
normalnego

(idea testu t-Studenta)

# Testowanie hipotez statystycznych

- Stawiamy hipotezę zerową (null hypothesis)
- Test jest procedurą obliczającą p-wartość, czyli wartość pewnej statystyki, będącej prawdopodobieństwem zaobserwowania danych przy założeniu prawdziwości hipotezy zerowej
- Jeśli p-wartość jest mniejsza od zakładanej, hipoteza zerowa jest odrzucana
- Testy statystyczne są obecne w każdym szanującym się pakiecie naukowym

# Testowanie metod optymalizacji

- Porównanie wielu algorytmów dla jednego zadania

	A1	A2	A3	A4
A1	.	+	+	-
A2	-	.	-	.
A3	-	+	.	-
A4	+	.	+	.

Na podstawie porównań w parach

	bilans	ranga
A1	1	2
A2	-2	4
A3	-1	3
A4	2	1

Agregacja tabeli porównań

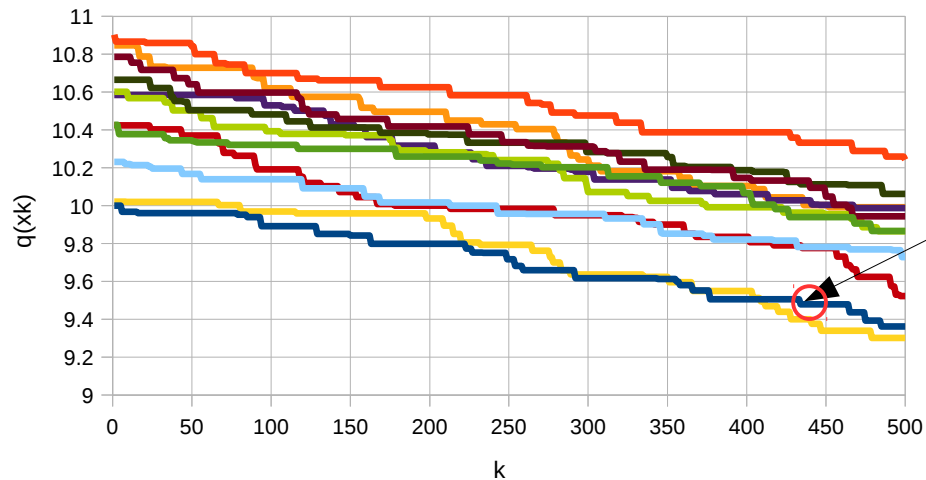
- Porównanie wielu algorytmów dla wielu zadań

	Z1	Z2	Z3	Z4	Z5	średnia
A1	2	3	1	1	4	2.2
A2	4	2	1	2	1	2
A3	3	1	1	2	2	1.8
A4	1	1	1	3	3	1.8

Wygrywają *ex aequo* metody A3 i A4

# Testowanie metod optymalizacji

- Porównanie wielu algorytmów dla jednego zadania – *function target pairs*

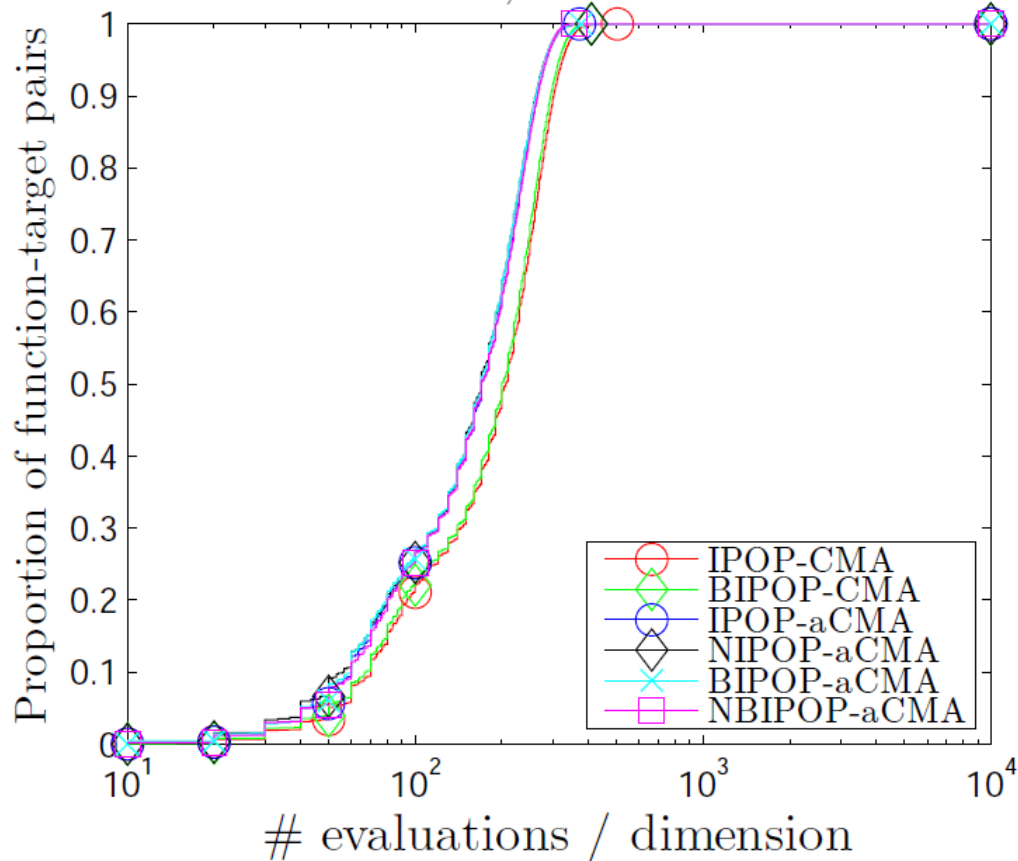


*function-target pair*

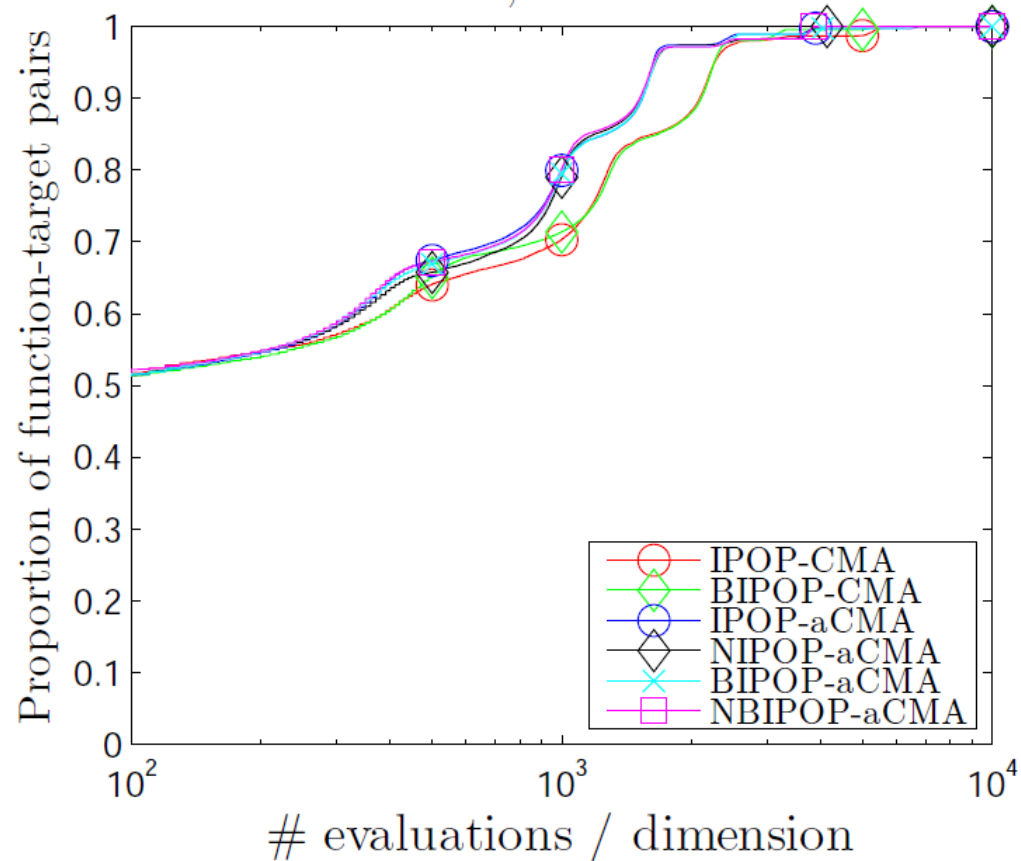
# Testowanie metod optymalizacji

- Porównanie wielu algorytmów dla jednego zadania: *function-target pairs*

F5 in 10-D, 30-D and 50-D

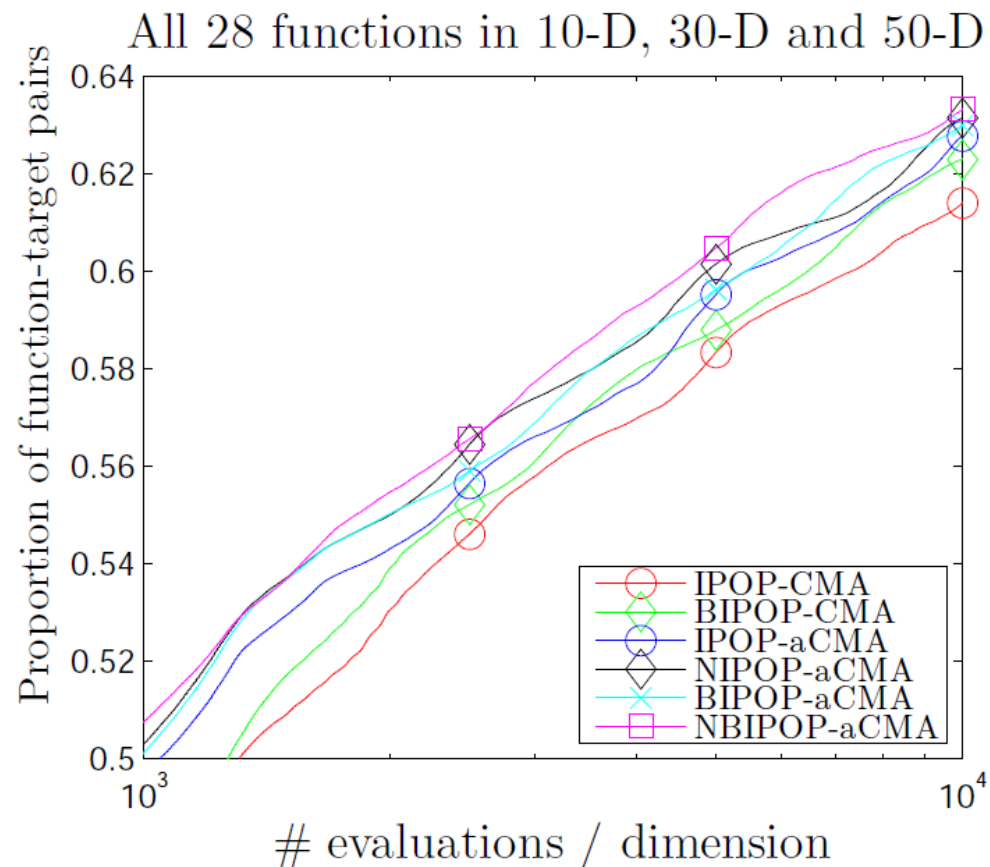
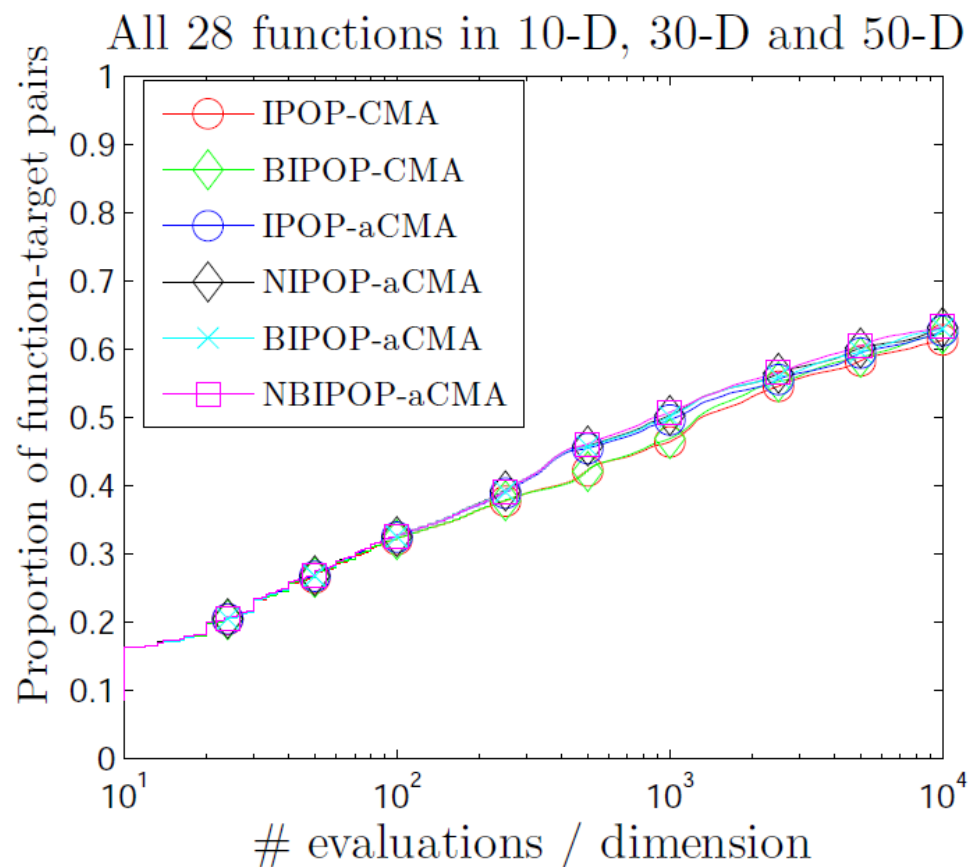


F6 in 10-D, 30-D and 50-D



# Testowanie metod optymalizacji

- Porównanie wielu algorytmów dla wielu zadań  
zadania: *function-target pairs*





# Testowanie hipotez statystycznych



Figures often beguile me, particularly when I have the arranging of them myself; in which case the remark attributed to Disraeli would often apply with justice and force: "**There are three kinds of lies: lies, damned lies and statistics.**"

- *Mark Twain's Own Autobiography: The Chapters from the North American Review*

# Effects of accepting a free lunch

